

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAAA
TTAAACAAGGCTGGCCTTGTACTTGAATACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCTTTTTGTGTCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTTCAACCCTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIYSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
 GTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGAGGCAGGAGCCTTC
 CTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTG
 GATTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAG
 GTGATCTTCTCCGTGACGTTTGCATTTTCTTGCAACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG
 AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGTGATCCTGG
 TTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACACGA
 CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCC
 CATTCTCAGCCCCAAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAG
 TGACTCTCATGGCTCTTCTTCTGGATTGGTGTCTCAACTGCCCATACACTTACATGTCTTACTTC
 CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
 CATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACA
 AACCATCAGGTTTCTGGGAATGATAAAAAGTGTACCACCTTCAGCATCAGGAAGTGAATCTTACT
 CTTATTCAACAGGAAGTGGATGCTTGGGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCT
 ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTTAATTTCTTGGTT
 ACTTTTTCTTATTTACTGTGTTTGAAAATTTTCATGGCTACCATCAATATTGTTTTGATCGAGTT
 GGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT
 GAAGTTTTGGTCCCAACACATTTCTTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGAT
 TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCTG
 CTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT
 AGAATACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTG
 ATGTGATCTTCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA
 GAGAAGCAAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAAATTTA
 GATATAAGAGGGGGGAAAAATGGAACCAAGGCCTGACATTTTATAAACAAACAAAATGCTATGGTAGC
 ATTTTTCACCTTCATAGCATACTCCTTCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
 AACATGAGAGGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
 TGTAGAGGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT
 CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTAAAGGTTACATGGAAAAGGTTATAGCTTTG
 CCTTGAGATTGACTCATTTAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCG
 ACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFEMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGBP
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDAAEELSRQ
LFLETADLYATKERIEYSKTFKGKYNFLGYFFSIYCVWKIFMATINIVFDRVGKTDVPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGMY
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
 GTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCATCCAGGT
 GTCATGCAGAATTATGGGGATCACCTTGTGAGCAAAAAGGCGAACCAGCAGTGAATTTACAG
 AAGCTAAGGAGGCCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAACTTGAGCTATGGCTGGGTGGAGATGGATTCTGGTCTATCTAG
 GATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTGCTGATTTGGAGGTTCCAGTGA
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCATTCCAGAA
 ATTATCACCACCAAAGATCCCATATTCAACACTCAAACCTGCAACACAAACAACAGAAATTTATGT
 CAGTGACAGTACCTACTCGGTGGCATCCCCCTTACTCTACAATACCTGCCCCCTACTACTCCTC
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAAATTGATTGTGTACAGAAAGTTTTTATG
 GAACTAGCACCATGTCTACAGAACTGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGA
 AGCTGTGGGTTTGGAGGTGTCCACGGCTCTGCTAGTGCTTGTCTCCTCTCTTTGGTGTG
 CAGCTGGTCTTGGATTTTGCTATGTCAAAAGGTATGTGAAGGCCCTTCCCTTTTACAAACAAGAA
 CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCTAA
 TGAGGAATCAAAGAAAAGCTGATAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGC
 GATGCCCTGGAAGCTGAAGTTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTT
 CATGCTCCTTACCCTGCCCCAGCTGGGGAAATCAAAGGGCCAAAGAACCAGAAAGAAAGTCCA
 CCCTTGGTTCCCTAAGTGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAGAGAAT
 GCCCTTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCACGGCCT
 TTCTAGCCTGGCTATGTCTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAAGTGCAAGGAC
 CTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCCTCCTGGCTGTCTGAGGCTAGGTGGGTG
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCGCGAGCTCAGACCCCTTCTTCA
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAGAGCAAAAGAAT
 GGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTCTCAGCAGGGACTGTAAAC
 ACAGACAGGGTCAAAGTGTCTTCTCTGAACACATTGAGTTGGAACTCACTGTTTAGAACACACACA
 CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA
 AAAATAAAAACCTTTATAAATTTCTATTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCACAAACATTTGCTGAATAGCTACTATATGTC
 AAGTGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATGACATAGTAG
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAAT
 TTTTATTTTGTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATT
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTTTAAAGTGCC
 ATTAACAAATGTATCACTAGCCCTCCTTTTTTCCAACAAGAAGGACTGAGAGATGCAGAAATATT
 TGTGACAAAAAATTAAAGCATTGAGAAACTT

FIGURE 6

MARCFSLVLLLSIWTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTAKEACRLLG
 LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLWKVPVSRQFAAYCYN
 SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPAPASTSIPR
 RKKLICVTEVFMETSTMSTETEPFVENKAAFNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
 RYVKAFPFTNKNQOKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCGCTCCCGCATCTGCACCCGAGCCCGGC
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCCGGGCGG
CGGCTGCGGGCGCAGAGCGGAGATGAGCGGGCTTGGGGCCACCCTGCTGTGCCGTGCTGTGGCGG
CGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG
GCTCTCAGCTACCCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAAGTATGAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACACAGAC
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAAGTGTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGC
ACACCCCTGCCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCCAGAGAGGTCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCC
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTA
CATCTTCTTCCAGTAAGTTTCCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTACAGT
CCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTCTTGGGAGAGTCAGGCAGGGTTAACTGCA
GGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCCTTACCAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTTGATAATTGTTGAGGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAAACATCAACCTGGCAAAAATG
CAACAAATGAATTTTCCACGCAGTTCTTCCATGGGCATAGGTAAGCTGTGCCCTTCAGCTGTTGC
AGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCT
CTCCACTACCCACACACAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACCTAATCTCACATCCCTCTAAAAGTAAA
CTACTGTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACCTACGCCACTCCACAAATGATG
TTTTCAGGTGTGATGAGTGTGGCACCATTGATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTAAATATGTATAAACATAAGTTGCATTTAGAA
ATCAAGCATAAATCACTTCACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLANLPPSYHNETNTDTKVGNNITHVHREIHKITNNQTGMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQGRMLCTRDESECCGDLVWGH
TKMATRGSNGTICDNQRDCQPLCCAFQRGLLFPVCTPLPVEGELCHDPASRLDLITWELEPDG
ALDRCPASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAAALLGGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
 GGCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCAGGGCTACTCATCCAAAG
 GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTCTGGACCCTT
 AACTGGGTACTGGCCCTGGGCCAATGCGTCCCTCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT
 CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
 AACTGGGTACTGGGCATTTGGAGCCCTCATCTGACCCTTGTGCAGATAGCCCGGTGCATCTTG
 GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCCGCTGCATCATGTGCTGTTT
 CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACGCAATGCATACATCATGA
 TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC
 ATTGTCAGGGTGCTCGTCCCTGGACAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
 GGTCGGAGGCGTGGGGTCTGTCTTCTTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG
 ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCTAT
 GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCTCTGCTTCCT
 GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA
 AGATTCTGGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
 CCCTGATCCAGGACTGCACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
 CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACT
 TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG
 AAACCTCCGTCTCTATTAATAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCTATCCCA
 GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA
 GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAA
 AAGATTTTATTAAAGATATTTTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGFWTL
 NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
 EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
 IVRVVLDKVTDLLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
 VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:**Transmembrane domains:**

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLI FTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFILIQLVLLIDFAHSWNQRWLKAE
ECDSRAWYAGLFFF TLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGVYTYSY
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVVKICASWAGLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA
 GGTGAAAAAGACTCCTGTAACCCCTCCTCCAGGATGAACCACTGCCAGAAGACATGGAGAACG
 CTCTACCCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
 CTCATGGCCAGGATTGAGTCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
 TTTCTGTTTGTGTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG
 TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
 TTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG
 ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG
 TGATCCTTTTGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTTCATTATC
 CTTGCCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
 CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
 GTCAGTTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
 GAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTAAATGTGAAAAACCTCACAGAAAGTC
 ATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC
 CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC
 CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT
 TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTTGTGCTGTATTCCTAATC
 AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTTATTTCAATTCAC
 AGAATGGAATTTTTTGTTCATGTCTCAGATTTATTTTGTATTCTTTTTTAACACTCTACATT
 TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTAAAAAGTGAATAAAATCTG
 ACATGTCAATGTGGCTAGTTTTATTTTCTTGTGTTTGCATTATGTGTATGGCCTGAAGTGTGGGA
 CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAATGTCACCAGACATTTGTATTATTT
 TTATCATGAAATCATGTTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTGAATGC
 ACAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTTCAGCCAATTTCAATTAATGAA
 CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:**Signal peptide:**

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCCGACCCGCCAGGAAAGACTGAGG
 CCGCGGCTGCCCCGCCCGCTCCCTGCGCCGCGCGCCCTCCCGGGACAGAAGATGCTCCAG
 GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCTGGGGTGACGGGCTGCCCAT
 CCGGCTGCCAGTGACGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACACCGGTGCC
 CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
 AGGCAGCTTTGCCGCGCTGCCGGGCTGCAGCTCCTGGACCTGTCACAGAACCAGATCGCCAGCC
 TGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
 CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTTACCTGGGCAAGAA
 CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC
 AGGACAACGAGCTGCGGGCACTGCCCCGCTGCGCCTGCCCCGCTGCTGCTGCTGGACCTCAGC
 CACAACAGCCTCCTGGCCCTGGAGCCCGCATCCTGGACACTGCCAACGTGGAGGCGCTGCGGCT
 GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTACGCGCTTGCGCAACCTCCACGACC
 TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCCCTCCGGGGCTGACG
 CGCCTGCGGCTGGCCGGCAACACCCGCATTGCCAGCTGCGGCCCCGAGGACCTGGCCGGCCTGGC
 TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC
 TCTTCCCCCGCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGC
 TGGTTTGGCCCCCTGGGTGCGCGAGAGCCACGTCACTGAGGAGCCCTGAGGAGACGCGCTGCCA
 CTTCCCGCCCAAGAAGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCGACTTTGGCTGCCCAG
 CCACCACCACACAGCCACAGTGCCACCACGAGGCCCGTGGTGGGGAGCCACAGCCTTGTCT
 TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTC
 CACTGCCCCACCGACTGTAGGGCTGTCCCCCAGCCCCAGGACTGCCACCGCTCCACCTGCCTCA
 ATGGGGGCACATGCCACCTGGGGACACGGCACCTGGCGTGTGTGCCCCGAAGGCTTCACG
 GGCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCACGCCGAG
 GCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC
 AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG
 GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACAGGTACCCCA
 GCTGCGGCCCCAACGCCACTTACTCCGTCTGTGTCATGCCTTTGGGGCCCCGGCGGGTGGCCGAGG
 GCGAGGAGGCTGCGGGGAGGCCCATACACCCAGCCGTCCACTCCAACCACGCCCCAGTCACC
 CAGGCCCCGCGAGGGCAACCTGCCGCTCCTCATTTGCGCCCCGCTGGCCGCGGTGCTCCTGGCCG
 GCTGGCTGCGGTGGGGGACGCTTACTGTGTGCGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG
 ACAAGGGCAGGTGGGGCCAGGGGCTGGGCCCTGGAAGTGGAGGGAGTGAAGGTCCCTTGGAG
 CCAGGCCCCAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC
 ACTCATGGGCTTCCAGGGCCTGGCCTCCAGTCACCCCTCCACGCAAAGCCCTACATCTAAGCCA
 GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC
 ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGACAGACAGGGCTGTGTGACCCAGCT
 GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
 CTAACGTCCCCAGAACCAGAGTGCCATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
 CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCCTGGGTCTGCTGGGCTCTCCAC
 TCCAGGCGGACCCCTGGGGGCCAGTGAAGGAAGCTCCCGAAAGAGCAGAGGGAGAGCGGGTAGGC
 GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGC
 TTTAGGAACATGTTTGTCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCATTTATCT
 GGGAAGATGTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGAAGACAAACGATGATATGAA
 GGCTTTTGTGAAGAAAAATAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLPLLLLLLALGPGVQGCPSGCQCSQPQTTFCTARQGTTPRDPVPPDTVGLYVFENGIT
 MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
 LGKNRIRHIQPGAFDTLDRLELEKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGILDTANVE
 ALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
 AGLAALQELDVSNLSLQALPGDLSGLFPRLRLAAARNPFNCVCPLSWFGPWVRESHTLASPEE
 TRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPTWSPTAPATEAP
 SPPSTAPPTVGPVPQPDCCPSTCLNGGTCHLGRHHLACLCEGFTGLYCESQMGQGTRPSPTP
 VTPRPPRSLTLGIEPVSPSTSLRVGLQRYLQGSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEY
 TVTQLRPNATYSVCVMPGLGPRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNPLLIAPALAAV
 LLAALAAVGAAYCVRRGRAMAAAAQDKGQVGPAGPLELEGVKVPLEPGPKATEGGGEALPSGSE
 CEVPLMGFPGLQSPHAKFYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
 594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCCTCGGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGGAATGAAAATCCTTAATGGAAG
CAATAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTGAGAAGCTGACTGAGGAAGGCTCTCCAAGGGACAGACTGCTCTTGG
CTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGTATATTATACAT
TTGGAGCTCTTGGGGCAATCTAATAGCCACATGGTTTTGGTAAGTAGACTTTAGTGGAAGGCT
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTTTCAGCTTTCATGATC
CAGATTTGCTTGTATTAAGACCAATATTCAAGTTGAACCTCCTTCAAATCCTGTTAATGGATAT
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTCTTTAAAATGATTAG
TTTGGCTGATTGCCCCATAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTCTCTGAGTTG
GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA
AAATTTTGTAATGGTGGCTATAGAAAAACAACATGAAATATTATACAATATTTTGCAACAATGC
CCTAAGAATTGTTAAAATTCATGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTG
TTTTTTACTTTTCATGATTGGCTGTCTTCCATTATTCTGGTCATTTATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTGCTAATTTGG
AAGATTAACCTCATTTTAAATAAAATTATGTCTAAGATTAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLSKTTLTSDSVKDHTTAGRVVAGQIFLDSESEL
 ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFPLFLDK
 EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEAAKRRQMQEAMMYQTGMKILNGSNKKSQKR
 EAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFELKTEEGSPKGQTALGFLYASGLGVN
 SSQAKALVYYTFGALGGNLIAMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAGCCATTCTGCAGTGGAATTCATGAACTAGCAAGAGGACACCATCTTCTT
 GTATTATACAAGAAAGGAGTGACCTATCACACACAGGGGAAAAATGCTCTTTTGGGTGCTAGG
 CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAATAAGATTGAAGACATCACTG
 ATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAACTTTTGAT
 AAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
 AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
 CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGTT
 CCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
 CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG
 TTATTAATGTCTCCAGTGTTGGAGGTCGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA
 TATGCAGTGGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGCTC
 ATGCATTGAACCAGGATTGTTCAAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
 TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAACAACAATATGGAGAAGGTTACATTGAAAA
 AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG
 CATGGACCACGCTCTAACAAGTCTCTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAA
 TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA
 GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA
 AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT
 CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCAGGGT
 CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCTGCCCT
 GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA
 TCTTTACCGTGGCTGCCCCATGCTTATGGTCCCAGCATTTACAGTAACCTTGTAATGTTAAGT
 ATCATCTCTTATCTAAATATTAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
 STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
 REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
 AFGVHVSCIEPGLFKTNLADPVKVEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGNKSYVNMD
 LSPVVECMDHALTSLEPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAE LANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
 CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTCTTCTTGGGGAAGTAAAA
 GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
 CATTCAAGAAATATATTCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
 AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATCCGT
 CGTCATTCAAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCAAAAACTTGCAGGAGCATTT
 TTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC
 ATCGACTGGAACATTCTTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGGTT
 GCCAATCTGGGCATGCTGAACAACCTGGGTTATAAACTGTATCAGGTTCTGTATGTCCACTGG
 TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTTGAAGAAGATGGATCCTTAAAGGAGG
 TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAAGTG
 GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA
 GAAAAGGAGAGGAGCACAGATTCAAGCAGCAAGAGAGAAGAACATCCAAAAGACCCCTCAGGAGA
 ACATTTTTCTTTGTGTCAGGCATTACGGACCTTTTTTCCAAATTCTGAATTTCTTCATTATGTGTT
 ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACTACAACCACCATCTCGATGTAGT
 AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCAC
 AAATCATTAAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA
 GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
 GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC
 CTACATTTTGATTCCTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC
 ATTTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTCAAC
 TGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA
 GATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
 ACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
 TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
 GAGACCATGTCTATTAAAAAATAAAATGAAAAGCAAGAATAGCCTTATTTTCAAATATGGAAA
 GAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCCTTAAGTGATACTTTTTTAGAAGTA
 CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
 AAAATTTAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLGGEVKGEAKNSITDSQMDDVEVVYITIDIQKYI
 PCYQLFSFYNSSGEVNEQALKKILSNVKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
 VFLLLTPSIITESCSTHRLHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
 QTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
 QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDDVDNLTL
 MVEHTDIPEASPASTPQIIKHKALDLDLRWQFKRSRLDQDKRSKANTGSSNQDKASKMSSPET
 DEEIEKMKGFGEYSRSPTF

Important features:**Signal peptide:**

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGACGAGCGGACCAGCGCAGGGCAGCCCAA
GCAGCGCGCAGCGAACGCCCGCGCGCCACACCCCTCTGCGGTCCCCGCGCGCCTGCCACCCCTCCCTCCTTCCCC
GCGTCCCCGCTCGCCGCGCAGTCAGCTTGCCGGGTTGCTGCCCGCGGAAACCCGAGGTCACCAGCCCGCGCCTCT
GCTTCCCTGGGCGCGCGCCCTCCACGCCCTCCTTCTCCCTGGCCCGCGCCTGGCACCGGGGACCGTTGCCTGA
CGCGAGGCCAGCTCTACTTTTCGCCCGCGCTCTCTCCGCTGCTCGCCTCTTCCACCAACTCCAACCTCTTCTCCC
TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGCGCGCTTCCCGTCCGGTCCCAAA
GGTGGGAACGCTCGCCCCGCGCGCACCATGGCACGGTTCGGCTTGCCCGCGCTTCTCTGCACCCTGGCAGTGCTC
AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGACGCTTTACGTGTCCAAAGGCTTC
AACAGAAGCATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCAGGGTTTACCTGTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAAGTAAAGATGATTTCAAAGATGTGGTCAGCGAACAGTGCAATCATTTG
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTGATGAATTTCTCAAAGAACTACTTGAAGATGCAGAGAAATCCCTG
AATGATATGTTTGAAGACATATGGCAATTTATACATGCAAAATTTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTG
AAAGCTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG
TTCCGCTGGTGAAGTCCAGTACCCTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGAGCAGCTGAAG
CCCTTCGGAGATGTCCCTCGCAAATGAAGCTCAGGTTACTCGTGCTTTTGTAGCAGCCGCTACTTTCTGCTCAAGGC
TTAGCGGTTGCGGGAGATGCTGTGAGCAAGTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTGAAG
ATGATCTACTGCTCCCACTGCCGGGTCTCGTGACTGTGAAGCCATGTTACAACACTACTGCTCAAACATCATGAGAGGC
TGTTTGCCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCTTTCAACATTGAATCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTTCAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
TCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCGAGGAACGCCCAACACAGCAGCTGGCACT
AGTTTGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCTCCCTTCCGAGCAAC
GTTTGCAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC
CTGTTTGCACTGACAGGAAATGGATTAGCCAACAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACAGAC
ATACTGATCCTTCTGCAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGTGGTGTCCGCTCTGGGCA
CAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGGTTATGCAGAGAGAGTGGAGATTAATTTCTCAAACCTGAG
AAAAAGTGTTTCATCAAAAAGTAAAAGGCACCAAGTTATCACTTTTCTACCATCCTAGTGACTTTGCTTTTAAATGAA
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAGTGCTGACTTTGTTTTCTATTAGTTTTGGG
AGGAAAAGGGACTGTGCATTGAGTTGGTTCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGAGGTACAGAA
CTATAGTTAGTTGTGATTTGTGATTTTATCACTCTATTATTGTTGTATGTTTTTCTCATTTCGTTGTGGGTT
TTTTTTTCCAACTGTGATCTCGCTTGTTCCTTACAAGCAAACAGGGTCCCTTCTTGGCAGTAACATGTACGTATT
TCTGAAATATTAATAGCTGTACAGAAGCAGGTTTTATTATCATGTTATCTTATTAAGAAAAAGGCCAAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
 CCSQEMEKEYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSINDMFVKTYGH
 LYMQNSELFKDLFVELKRYVVGNVLEEMLNDFWARLLERMFRLVNSQYHFTDEYLECVSKYTE
 QLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVNPTAQCTHALLKMIYCSHCRGL
 VTVKPCYNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
 MQDNSVQVSQKVFQGCQPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
 EKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNPEVQVDT
 KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
 KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTGCCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGAATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGAGCTCTG
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCAÇACCTACCAGACACTC
TTCTTCTCCACCTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTTTTTCAAGATCATTTTGTTTGTGCTCTCTCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT
AGCTAGTGTCAATTTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPDHFKNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
 AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC
 CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCTACCTGA
 TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
 ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
 CAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
 AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
 GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTGTG
 CTCTGTGCAGAGCCCAAGGAAGGTCCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
 TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTC
 GAGCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
 GAACGCCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
 GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTCCCAAGCTCCAAGGCACTCATTTGCTCC
 TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTA
GCAGAATGAGAGAAGACATTTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
 AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
 AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC
 CCAATGTTGTCCCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACAC
 CCAATGCGTCTCTAGGAAGTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
 CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG
 ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTTCCTCAAATATTTTTTAATAATAGACGAA
 ACCACG

FIGURE 28

MDILVPLLQLLVLLLTLPPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSSQIKGL
 TGASGKVALLELGCGTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
 RQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQVFEPTW
 KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHINGKAVKQSFSSKALICSFPSL
 QLEQATHQPIYLPRLGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCTATCCACCTCCCCAAGCCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCCAGAAAGTCTCTTCTG
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAA
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTATTAAAGAAATT
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
 CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
 ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGAATCATGTCGG
 GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGCTTTGGCCATGATGTTTACC
 TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTATTTTGGGATTGTTGTT
 TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATAACCAACGACCTCAGCATAGAATTGGACA
 CAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
 CTGCTCGTCTTGATTTTGTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCAAATCAC
 AAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCAGCCACTGTGGACATTTGCCATCTCA
 TTTTCTTCTGGGTCCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAACTGCAGGAGCTGCCAGGTT
 ATGGAAGGCGGCCAAGTGAATATAAGCCCCCTTCGGGCATTCCGTACATGTGGTCGTACCATTT
 AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAG
 TGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCATCCTTTCTGCTCTC
 TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTAACTCTCTGTGGTGAG
 GATTCCGAGAATCATTTGTCATGTACATGCAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
 CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
 AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
 ATTCAAAATCTTGTCGAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA
 TTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC
 AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC
 CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCCTGTGTTTGTGTTGATC
 TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
 AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGA
 GGGAACAGAACTCCAGGCCATTGTGAGATAGATATCCCATTTAGGTATCTGTACCTGGAAAACATT
 TCCTTCTAAGAGCCATTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT
 TTAAGACCTAATAAACCCCTATTCTTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYTNDLSIE
 LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLFQPLWTFA
 ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
 GAVVTCYFNRSKNPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQNALKEQQHG
 ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
 FIIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
 VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:**Signal peptide:**

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCCTTAGA
 ATAATTTGTATGGGATTTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCATTCTGTGTGGT
 GAAAATTTTTTGAAGGAAATTCCTTCTTCAAACAAGGGTGTCAATCTGATATTTATGAGGAC
 TGTGTCTCTCACTATGAAGGCATCTGTTATTGAAATGTTCCCTGTTTTGCTGGTGACTGGAGTAC
 ATTCAAACAAGAAACGGCAAAGAAGATTAAAAGGCCAAGTTCAGTGTGCCTCAGATCAACTGC
 GATGTGAAAGCCGGAAGATCATCGATCCTGAGTTCATTGTGAAATGTCCAGCAGGATGCCAAGA
 CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG
 TACACAGTGGTGTGCTTGATAATTCAGGAGGGAAAATACTTGTTCGGAAGGTTGCTGGACAGTCT
 GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT
 TATCGTCTTAGAAAGTAAACCCAAAAAGGGTGTAACCTACCCATCAGCTCTTACATACTCATCAT
 CGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
 ACAACTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGTCACTGTAGCTGTGGCCACCCAC
 CACCTTGCCAAGGCCATCCCCCTTCTGCTGCTTCTACCACCAGCATCCCCAGACCACAATCAGTGG
 GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCC
 AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTTGGAGC
 GGATGTGAGCCTGGGACTTGTTCAAAAGAAGAATTGAGCACACAGTCTTTGGAGCCAGTATCCC
 TGGGAGATCCAACTGCAAAATTGACTTGTCTGTTTAAATTGATGGGAGCACCAGCATTGGCAAA
 CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCAAGCTCTTGACATTGGCCCTGC
 CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCCTGCTACTCACTTTAACCTCAAGACAC
 ACACGAATTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAAT
 GTAGTTCGGGCCATCTCCTTTGTGACCAAGAATCTTTTCCAAAGCCAATGGAAACAGAAGCGG
 GGCTCCCAATGTGGTGGTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
 GACTTGGGAGAGAGTCAGGAATCAACATTTTCTTCATCACCATTGAAGGTGCTGCTGAAAATGAG
 AAGCAGTATGTGGTGGAGCCCAACTTTGCAAAAGGCCGTGTGAGAACAAACGGCTTCTACTC
 GCTCCACGTGCAGAGCTGGTTTGGCCTCCACAAGACCCTGCAGCCTCTGGTGAAGCGGGTCTGCG
 ACACTGACCGCCTGGCCTGCAGCAAGACCTGCTTGAATCGGCTGACATTGGCTTCGTCATCGAC
 GGCTCCAGCAGTGTGGGACGGGCAACTTCCGCACCGTCTCCAGTTTGTGACCAACCTCACCAA
 AGAGTTTGAGATTTCCGACACGGACACGCGCATCGGGGCCGTGCAGTACACCTACGAACAGCGGC
 TGGAGTTTGGGTTCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
 TACTGGAGTGGTGGCACCAGCACGGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTCTTCAAGAA
 GTCCAAGCCCCAACAAGAGGAAGTTAATGATCCTCATCACCGACGGGAGGTCTACGACGACGTCC
 GGATCCAGCCATGGCTGCCATCTGAAGGGAGTGATCACCTATGCGATAGGCGTTGCTGGGCT
 GCCCAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCACTCCTTCTTTGTGGACGA
 GTTTGACAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTTGTACAGAGTTCAACTCAC
 AGCCTCGGAAC**TGA**ATTTCAGAGCAGGCAGAGCACCAGCAAGTGTGCTTTACTAACTGACGTGTT
 GGACCACCCACCCGCTTAATGGGGCACGCACGGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC
 AAATGTCTTGTTATTATCTTTGCCATCATGCTTTTTCATATTCAAAACCTTGAGATTACAAAGA
 TGATCACAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGAGAGATTTTACAT
 TTTGACAATTGTTTTCAAATAAATGTTTCGGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
 AGCTTTTGTGAGATTTTAAAGTTGTTATTTCTGATTTGAACTCTGTAACCTCAGCAAGTTTCAT
 TTTTGTATGACAATGTAGGAATTGCTGAATTAATGTTTGAAGGATGAAAAATAAAAAAAAAA
 AA
 AAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
 CQDPKYHVYGTDVYASYSSVCGAHVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWR
 ESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
 TPTTLPRPSPSAASTTSIPRPQSVGHRSEQMDLWSTATYTTSSQNRPRADPGIQRQDPSGAAFQKP
 VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
 GPAGPLMGVVQYGDNPATFNLKTHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
 RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENKQYVVEPNFANKAVCRTNG
 FYSLHVQSWFGLHKTQLVLRVCDTDLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
 LTKEFEISDTDTRIGAVQYTYEQRLFEFGFDKYSSKPDILNAIKRVGYWGGTSTGAAINFALEQL
 FKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAWAAQEELEVIATHPARDHSFF
 VDEFNLHQYVPRIIQNICTEFN SQPRN

Important features:**Signal peptide:**

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
 431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT
ACAGAAAACAACAAAAAATAAGCTTTAATTTTCATCTGGAATTCACAGTTTTCTTAGCTCCCTGGACCC
GGTTGACCTGTGGCTCTCCCGCTGGCTGCTCTATCACGTGGTGTCTCCGACTACTCACCCCGAGTGTA
AAGAACCTTCGGCTCGCGTGTCTGAGCTGCTGTGGATGGCTCGGCTCTCTGGACTGTCTTCCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTCACTCCTGAGTTCTTTGTGATGTGGTAC
CTCAGCCTTCCCCACTACAATGTGATAGAACCGGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTGA
CAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAACTGCTCTCATCAAAATCCATTTCTGGTCATTC
TGGTGACCTCCACCCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTACTTGGGGTGAAAAAAGTCT
TGGTGGGGATATGAGGTCTTACATTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC
ATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA
ACCTGACCTTGAAAACCATTTATGGCATTCAGGTGGGTAACTGAGTTTTGCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA
GAAGTTTTTTCACAGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCATATTT
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATATAATGTCCAGAGAT
TTGGTGCCAAAGGATCTATGAAATGATGGGTACGTAAAACCCATCAAGTTTGAAGATGTTTATGTCTGGGAT
CTGTTTGAATTTATTAAAAGTGAACATTCATATTCAGAGACACAAATCTTTCTTTCTATATAGAATCC
ATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTT
TGGCAGGTGATGCTAAGGAACACCACATGCCATTATTAACCTTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTTGTGGAAAGTGTAAATAAAGTAGGTACTGTGGAATAATTCATGGGGAGGTGAGTGTCTGGCTT
ACACTGAACTGAACTCATGAAAACCCAGACTGGAGACTGGAGGGTTACACTGTGATTATTAGTCAGG
CCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAAAGAAATTAATAGG
ACCAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGTAAAACAAAACAATGTAGAGTTTTATTTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA
TATCTTATGTGGATTACCAATTTAAAAATATATGATGTTCTGTGTCAAAAACCTCTTCACTGAAGTTATA
CTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT
ATTATTTAAAATTACTTCACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAG
TGAATCATCTTTTACATGCAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCACTCCA
TTAATGTAAAGTCATAGGTCAATTATTGCATATCAGTAATCTCTGGACTTTGTTAAATATTTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPiYRQDFHF
 TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EKEDKMLA
 LSLEDEHLLYGDIIRQDFLDTYNNLT LKTIMAFRWVTEFCPNAKYVMKTD TDVFINTGNLVKYL
 NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPYCSGLGYIMSRDLVPRIYEMMGHV
 KPIKFEDVYVGICLNLKVNIIHIPEDTNLFFLYRIHLDV CQLRRVIAAHGFSSKEIITFWQVMLR
 NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCCGCGCAAGGATGAGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCCTTCATC
TCGTCTTGGCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGAATATCATGTGTCGGGAGTGCTGTG
AATATGATCAGATTGAGTGCCTTCGCCCGGAAAGAGGGAAGTCGTGGGTTATACCATCCCTTGTGTCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAACGCAAGAGCTGCCGAAATGGCTCATGGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAGTGCAGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTTGTGGAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTCATGC
TAAACCTGGGTTTGTGCATCCAACCTAAGATTTGTGCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTCTGTATGGAGACAACCGCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCCTCACTCCAGCTCCTCTTCCACTCCGATGGCTCCAAGAATTTGACGGTTTCCATGCCATTATGAGGAGATCACAGC
ATGCTCCTCATCCCTTGTTCATGACGGCACGTGCGTCTTGACAAGGCTGGATCTACAAGTGTGCTGCTTGGCAGGC
TATACTGGGACGCGCTGTGAAAATCTCCTTGAAGAAAGAACTGCTCAGACCCTGGGGGCCAGTCAATGGGTACCAGAAAA
TAACAGGGGGCCTGGGCTTATCAACGGAGCGCATGCTAAATTTGGCACCCTGGTGTCTTTCTTTTGAACACTCCTATGT
TCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAAAGGAGAGTGGTCAGGGAACAGCCCATCTGCATAAAGCCTGCCGA
GAACCAAGATTTAGACCTGGTGAGAAGGAGAGTTCTTCGATGCAGGTTGAGTCAAGGGAGACACCATACACAGCTAT
ACTCAGCGGCTTCAGCAAGCAGAACTGCAGAGTGGCCCTACCAAGAGCCAGCCCTTCCCTTTGGAGATCTGCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCTTCTACCCGCGCTGGGACGAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGGCGGGCACCACTCCGATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGA
CCCAAGGGTTGCGCTGGCGGTGGCAGGCAGCCATCTACAGGAGGACAGCGGGGTGCATGACGGCAGCTACACAAGGGAGC
GTGGTTCTTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCATGTGTTACTGACCTGGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTTGGGGAAATCTACCGGATGATGACCGGGATGAGAAGACCATCC
AGAGCTACAGATTTCTGCTATCATCTGCATCCCACTATGACCCCATCTGCTTGTGATGCTGACATGCCATCCTGAAGCT
CCTAGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCCTCGCTGCCAGTCCGGGATCTCAGCACTTCTTCCAGGAG
TCCACATCACTGTGGCTGGCTGGAATGTCTTGGCAGAGCTGAGGAGCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCCCTTCTGATATCTGCAGTGCAGAGACAGGAGGCATCGCGGCTGTGCTTCCCGGA
CGAGCATCTCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAACATGCAGCCACAGGCTCTCCA
CTGCTTACCAAGGTGCTGCTTTTAAAGACTGGATTGAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTTCTGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTGGGCTGTGAAGTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGTTGAGTAGACCTCCATTGCTGGTGGTGTGTCGCGCTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAGAAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGCTTCCCAACTTTTCAAGTATACGAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAG
GCCCCCTTTGAGGCTCTCAAGTCTTAGAGAGCTGCCTGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCTT
TGTTACATGGGCACAGTACAGTCTGGTCTTTTCTTCCCATCTCTTGTACACATTTAATAAAATAAGGGTTGGCTTCT
GAACACAAAAA
AA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT
 IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGFYCAECRAGWYGGDCMRCGQ
 VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGOII
 KRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
 ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
 KRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPQVQSRETPHQLYSAAFSKQKLQSAPTK
 KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIENITAP
 KTQGLRWPQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVAAHCVTDLGKVTMIKTADL
 KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
 DLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
 SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMLVSWSYDKTCSHRLSTAFTKVLFPKDWI
 ERNMK

Important features of the protein:**Signal peptide:**

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
 474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATTCCTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAAATGAAGGCGCTTCTGTTGC
TGGTCTTGCCTTGGCTCAGTCTGCTAACTACATTGACAATGTGGGCAACCTGCACCTTCTGTATTAGAACTCTGTA
AAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCAAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCCTGGCCTAGACA
ACCTTGCTACGTGCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCAGTGGACTTGGCCGGAGCAACCGAACTA
GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTCTTCGAA
GGACAAAGAGCGGGAGTGAGTTGCCAACCATGCCGACCAGGGCAGGGGAAAAATCTGAAAACACCACTGCCCTGAAG
TCTTTCCAAGGTTGTACACCTGATTCCAGATGGTGAATTAACAGCATCAAGATCAATCGAGTAGATCCCACTGAAA
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTTATCCAACACATTTATCGTGATGGGG
TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACTACGCTGTGCGTCTCCTGCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTCCGCA
GCAGGAACAATGGACAGGCCCCGGATGCCACAGACCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTTGAATAAACTGGTGGCAAGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCACTTGGAGAGATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCA
GCCCAGAAAGTGGGCTCATCTGATTGAGGCCAGTGAAGACGTGTTACCTCGTGTGTCGCCCGAGGTTGGGCAGC
GGAGCCCTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCAGGGCCAGGGGAGAGGAGCAACA
CTCCCAAGCCCCCTCATCTACAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCCGTGAATCTCTCG
GCATGACCGTCGCAGGGGGAGCATCACATAGAGAATGGGATTTGCCATCTATGTGATCAGTGTGAGCCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAACAGCAGAGGTGAGCC
GGAGTGAGGCAGTGGCATTTATTGAAAAGAATCATCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGAAGTGGTCCCCATCCT
GGGTCATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACAGCTGGAAGTC
TGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAAACAAACCTTTTTTCATCAAATCCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTGCTGTCAATGGTAGAAGTACATCAGGAATGA
TACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT
TTTTATAGAATCAATGATGGGTGAGGAGAAACAGAAAAATCACAATAGGCTAAGAAGTTGAAACACTATATTATC
TTGTGAGTTTTATATTTAAAGAAAGAAATACATTGTAAAAATGTGAGGAAAGTATGATCATCTAATGAAAGCCAGTT
ACACCTCAGAAAATATGATTCAAAAAAATTAAGTACTAGTTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTAC
AACATTGTTTATATTTTTCTATTCAATAAAAAGCCCTAAACAACTAAAATGATTGATTTGTATACCCCACTGAATT
CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGCTACATTATGGCCATTTTAAATTTACAGCT
AAAAATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATTTTTTCAGAAGTTAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS
 PEVSAAATISLMTDEPLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINR
 ALSVLRRTKSGSAVANHADQGRESENTTAPVFPRLYHLIPDGEITSIKINRVDPSSELSIRLV
 GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
 REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFIENVLDGGVAYRHG
 QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQVRSPDIFQEAGWNSNGSWSPG
 PGRSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPYVISVEPGGVISRDGR
 IKTGDILLNVVGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
 SDWSPSWVMWLELPCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
 RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:**Signal peptide:**

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
 467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
 CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACAGTAATAAAAAACATGGGCTTCAACCTGACT
 TTCCACCTTTCCCTACAAATTCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGTTGGGTGGGC
 CACCAGTAACACTTTCGTGGGTGCCATTCAAGAGATTCCCTAAAGCAAAGGAGTTCATGGCTAATTTCC
 ATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAAGTT
 GACAACTGTCCTTCTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACAGATCTCAC
 TTTGGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG
 CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
 CATCTGCATCCCTTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTATCCACCAGGCTGAAGG
 TAAAAAGTTTAAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
 ACTGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
 CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGG
 GGGTGTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACACTACTGGGGAT
 GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGGCCCTG
 CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACC
 GATGAAGCTCTTACACCAAGTGTACAGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT
 TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTTCTGGTTTGGTGCATTGACCC
 TGGATCTTTTGGTGATGTTTGAAGAAGTGAATCTTTGTTTGAATAATTTTGGCCTAGAGACTTCAA
 ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTCTCTTTTGTATTTTCT
 TAGCAGAGCTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGAT
 CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
 AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGAT
 AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
 AGGTACGAAGATACAATACTGTTATTTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTCAGGT
 GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA
 GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
 CTTCAGGGGAGGACCTGCCAGGTATGCCTTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT
 TTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
 ATTAACATAATAAATATGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

FIGURE 42

MGFNLT FHLSYKFRLLLLLTCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKGKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSYFGGV TALSREQFFKVNGFSNNYWG WGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWR TDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:**Signal peptide:**

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCACAACAG
ACGGGACAACCTTGCAGAGCTGCAACCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCTCTCCC
TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAATGGCTGGTTCTTTGTTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLSLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTTGTCAACATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAAGTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTGAGCAGAAACT
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCTGGG
AAACTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTFLVTIQP
 EGGTIIVTQNRNRERVDFPDGGYSLKLSKLKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
 PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
 VARNPVSRNFSSPILARKLCEGAADDPDSSMVLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE
 EKKRVDICRETPNICPHSGENTYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLTTPDPT
 PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
 291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA
CTTGCACTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTTAGGTCTATTGCTTGTGGGAATTCGGAGGTCTGTTTGGGCT
CAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
 TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCM LISIQALLKGPLMCNSPSNSNANCEFSL
 KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
 LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:**Transmembrane domains:**

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCGCCCCAGTGCCCTCTCCCCCTGCAGCCCTGCCCCTCGAACTGTGA
CATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGAAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCCTCATCACTCCAGGCTCTGCCA
CTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQQQHSPVPEKAIPILITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
 AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
 AGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
 GGGCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAG
 CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA
 GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGAGGGGCAGCTGGCTCTAAAGTCAGTGAG
 GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
 AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACTGGGCACGAGA
 TTGGCAGACAGGCAGAAGATGTCATTTCGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG
 GTGCCCTGGCCACAGTGGTGCTTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
 CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG
 GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
 GGGCCACCAAACTTTGGGACCAACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCACTGAG
 AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
 ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
 AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
 TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
 CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA
 CATAAACCCGGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGGAGCGGGGAATCTGGGATTACGGG
 CTTAGAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTG
 GAGGCTCTGGAGACAATTATCGGGGGCAAGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
 GGTGGAGTCAATACTGTGAACTCTGAGACGCTCCTGGGATGTTAACTTTGACACTTTCTGGAA
 GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC
 GCATCCCGTGACTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCACACTCCCTCCTTAA
 AACACCACCTCTCATCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA
 AA
 AA

FIGURE 52

MKFQGPLACLLALCLGSCEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
 EALQGQTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGV
 GHSGAWETSGGHGIFGSQGGLGGQGNPGGLGTPWVHGYPGNSAGSFGMNPQGAPWGGGNGGPPNF
 GTNTQGAVAQPGYGSVRASNQNEGCTNPPSGSGGGSSNSGGGSGSQSGSSGSGSNGDNNNGSSSGGS
 SSGSSSGSSSGSGSSSGSGSSGSGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKPGCEKPGNE
 ARGSGESGIQGRGQGVSSNMREISKEGNRLGGSGDNYRGQSSWGSGGDAVGGVNTVNSETSPGM
 FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
 90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
 159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
 236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
 253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
 283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
 298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
 325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
 CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
 ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAAGTCCGCCGGCTCCAGTGTTC
 CACAGCCCCAAAACGGAAGTGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC
 TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGTATGGCTGGGTCCCAT
 CATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
 TTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
 AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT
 GAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
 CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
 CAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
 CTTGGAGCTCAGTGGCCTTG TAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
 ATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC
 GCTGTCATCCGGGAGCGGCGTCGACCCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA
 AGCCAAGTCCAAGACTTTGGATTTCAATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
 CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCAGG
 GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
 ACAGGAGGTGCAAGAGCTTCTGAAGGACCGGATCCTAAAGAGATTGAATGGGACGACCTGGCCC
 AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATC
 TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG
 CCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
 CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCTTTCTCCGCA
 GGGCCCAGGAAGTGCATCGGGCAGGCGTTCCGCATGGCGGAGATGAAAGTGGTCTGGCGTTGAT
 GCTGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAATTGATCATGC
 GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCAGTTGACTTTCTGAC
 CCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGTGTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFFQPPKRNWFWGHLG
 LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIYVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
 WLGEIGILLSGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
 SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
 RLVHDFDPAVIRERRRTLPTQGIDFFKDKAKSKTLDLFDVLLLSKDEDEGKALSDEDIRAEADTF
 MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
 LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
 LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
 VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT
TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTGCAGTCTTTGGCATTGACGTGG
TACAGCCTTTCC¹TCATAACCATTGCAAGGGATGCTGTGAAGAAGTGTTTGGCCGTGTGCTTGC
ATAATTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAGCTGGTGGACAGTTTT
GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTATCTTGCAGCAATGTGTTGCTT
GTGATTCGAACATTTGAGGGTTACTTTTGGGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAACTGGGTTCGTATCTTGTGGAGTGAATCTTCCTCATGTACCTGTTTCCTC
TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTCAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCAGCAGGCCGCCAAT**ATG**CTCTGTCTGTGCCTG
TACGTGCCGGTCATCGGGGAAGCCAGACCGAGTTCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC
ATTTTCAAGCTCAGTGTCTTCATCCCTCCAGGAATTCACACCTACCGCCAGTGGGAAGCAGAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGGCGAGCTAGACTTTGAAGAATTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTTTAAGATTTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTGCCGGACTTG
GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCACAGAGCATGGATAAAAACGGCAGATGACCATCGACTGG
AACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAACATCCCCGAGATCATCTCTACTGGAAGCATTCCACG
ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGCAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGAGGTGGGGCAGGGCCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTGAGAAGGAGGGGCCAGG
TCACTCTGGCGGGGCAATGGCATCAACGTCTCAAAATTGCCCGGAATCAGCCATCAAAATTCATGGCCATGAGCAG
ATCAAGCGCCTTGTGGTACTGACCAGGAGACTCTGAGGATTACAGAGAGGCTTGTGGCAGGGTCTTGGCAGGGGCC
ATCGCCAGAGCAGCATCTACCAATGGAGGTCTGAAGACCCGGATGGCGTGGCGAAGACAGGCCAGTACTCAGGA
ATCTGGAGTGGCCAGGAGGATCTGGCCAGAGAGGGGTGGCCGCTTCTACAAAGGCTATGTCCCCAACATGCTG
GGCATCATCCCTATGCCGGCATCGACCTTGACGTCTACGAGACGCTCAAGAATGCCTGGCTGCAGCAGTATGCAGTG
AACAGCGCGGACCCCGCGTGTGTGTCTCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGCTAC
CCCTTGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTC
TTCAACATATCTCTCGGACCGAGGGGGCTTCGGGCTGTACAGGGGGTGGCCCCCAACTTCATGAAGGTTCATCCCA
GGCTGTGAGCATCAGTACGTGGTCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTCGCGGT**TGAC**GGGGGGAGGGC
CGCCCGGCAGTGGACTCGTGATCTCTGGGCGCAGCCTGGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACACT
AAGCTGTCTCGAGCCAAGCTGTGAAACCCTAGACGCACCCGAGGGAGGGTGGGGAGAGCTGGCAGGGCCAGGGCTT
GTCTGTCTGACCCAGCAGACCCCTCTGTGGTTCCAGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTCAGCAGG
CTCCGGGCTCACATGTGTAAGGACAGGACATTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCTTGGAGCGCGGCT
TAGTCTTCCATTTACCCCTTGACGCCAGCTGTGGGCCACGGCCCTGCCCTCTGGTGTGCCGTGCATCTCCCTGTGC
CCTCTGTCTGCCTGCCTGTCTGCTGAGGTAAAGTGGGAGGAGGGCTACAGCCACATCCACCCCTCTGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCCTCCAGGCCCTGACTTCCCAACCTACAGCATTGACGCCAATTTGGC
TGTGAAGGAAGAGGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATG
CTTGGGAGTGCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTGCTCTGAGCT
GGCTTGGACCTGTGAGGATGGGCCCCACCTCAGAACCAAACTCACTGTCCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTTGAGGGCGAAGGGCAGAGCGTTTGTGTCTCTGGGAGGGAAGGAAAGGTGTTGGAGGCCTTAATTATGG
ACTGTTGGGAAAGGGTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCCTGTCCAACCCAGCAGGGGCGCAGC
GGGACCAGCCCCACATTCACCTTGTGTCACTGCTTGGAACTTATTTATTTGTATTTATTTGAACAGAGTTATGTCTCT
AACTATTTTATAGATTTGTTAATTAATAGCTTGTTCATTTTCAAGTTCATTTTTTATTCATATTTATGTTTCATGGTT
GATTGTACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGGAGAAGGGGGCTTGGGGCGCTGCAGTTCACATCT
GTCCAGAGAAATTCCTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAAGGCAGCAGCCCTGGCTCCTTCTCTTGGCAG
GTTGGGGAAGGGCTTGGCCCCAGCCTTAGGATTTACAGGTTTACTGGGGCGTGGAGAGAGGGAGGAACCTCAAT
AACCTTGAAGGTGGAATCCAGTTATTTCCCTGCGCTGCGAGGGTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAG
TGAGGTGCCTCTCACTGTGAATTTGTGGTGGGCGGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCGTCCCTCCAGC
CTTCTGCTGCCCTTGCTTAACAATGCCGGCCAACCTGGCGACCTCACGGTGCATCTCCATTCCACCAGAATGACCTGA
TGAGGAAATCTTCAATAGGATGCAAGATCAATGCAAAAATTGTTATATATGAACATATAACTGGAGTCGTCAAAAAG
CAAAATTAAGAAAGAAATGGACGTTAGAAGTTGTCAATTAAGCAGCCTTCTAATAAAGTTGTTCAAAGCTGAAAAAA
AA

FIGURE 58

MLCCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDLDG
 QLDFEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEI MQSLRDLGVKISEQQA EKILKSMDKNG
 TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG
 AGAVSRTCTAPLDRLKVLMMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIAK
 FMAYEQIKRLVGSDQETLRIHERLVAGSLAGATAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
 ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNWLQHYAVNSADPGVFVLLACGTMSSTC
 GQLASYPLALVRTRMQAQAASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
 VYENLKITLGVQSR

Important features:**Signal peptide:**

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
 TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
 CACTCATCATTGGCTTTGGTATTTAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
 GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACCTTTTGAACCTGACATCAAACCTTTCTGATAT
 CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
 AGCTGTCGGAGCAGGATGAAATGTTTACAGAGCCGGACAGCAGTGTGTTGCTGATCAAGTGATAGTT
 GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
 CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
 AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC
 CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTTCTCGGAAGTCTCCAATAC
 CAGCTTTGAGCTGAACCTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA
 TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
 ACAGAAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGTAAACTCAAAGGCTTCTCTGTGTGT
 CTCTTCTTTCTTTGCCATCAGCTGGGCACCTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT
 GTGCCTTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTACAC
 CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
 AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
 CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTATGTGAACTAGACAAGTGTGTTAAGA
 GTGATAAGTAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
 CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTATGTTATATGTGC
 TGTAATGTTGCTCTGAGGAAGCCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATCCAC
 AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
 GCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC
 TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT
 CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTAAACAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
 DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLKKNVQLTDAGTYKC
 YIIITSKGKGANANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVWVASQVDQGAFSEVS
 NTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTESEIKRRSHLQLLNSKASL
 CVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
 220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCCTGGTGGTGGTTATGGGGGTCTGCCCCTGGAGGGCCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCCTCTGGAACCTCAGGAGGACCATATG
GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGGCCAGCAGCCT
GGGCTTTATGGACAGGGTGGCGCCCCCTCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
GGCCGCATCGATGCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCCAGTTCACCCAGCTTCTGGTCTCCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
CCATGACAGCTTCTCGGATGCTATTACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCTGCATCATAGCCACCA
AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAATGG
AGTTAGTGTCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGGAGTGTTCATCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAAT
CCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTAATTTGCATTTTTTTTC
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDP EAYSWFQSVDS DH
SGYISMKELKQALVNCNWSSFND ETC LMMINMF DKT KSGRIDVYGFSALWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGGGACC
AACTTGCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGAGCCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGGGA
AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGGCAGTGCCCTT
TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCCGTCTAATTCACTGGGAGCCGGGGGAAAAATCCTTTCCCAACG
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTG
TGCTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA
ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG
GCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCTCTCTGGAGTTCTCCGC
CCTCCTGGCTCTTCTTGGAAACATCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGGCTAGAGCACGATAGAGGGAACCCAACATTGGGAGTTAGAGTCTGCTCCCGCCCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGSPNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSQVQRWPPSWGLPAMDSWPPEDFWQMAAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSGWNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPFPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
 TGGGCTGCCCCCTTGTCTCCTCTTGACCCCTCCTGGCAGCTCACATGGAACAGGGCCGGGTATGA
 CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
 TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC
 AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCTTCTTGGCCC
 GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCCACCTC
 CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MGSGPLVLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGGGC
 CAGGTGCCCCGTCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGGAGAAGCCCC
 TTCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG
 CTGGGCCTGCCGTTCTTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC
 AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
 AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG
 GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCAGTAGCGA
 GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
 GCCTGCCCATCTAGGTCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
 GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA
 AGACTCTGCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACCTTTATATATTTATATAAAATTAG
 TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSTSSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLEKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTTCAGAGGTTTTGTTTTCTCTTAGTCTCTGTCCTGCTGCACCAG
 TCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATATTGTCATTGTTATAGATCCTAGTGTGCCAGAA
 GATGAAAAAATAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGAA
 TTTTTTTTCAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAAACATGAA
 AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCAGGTAGAGATGAACCATACACCAAGCAGTTTACA
 GAATGTGGAGAGAAAAGGCAATACATTCACTTCACCCCTGACCTTCTACTTGGAAAAAACAAATGAATATGGACCA
 CCAGGCAAACGTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGTTGATGAGTACAATGAAGATCAGCCTTTC
 TACCGTGCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT
 CAAGGAGGCAGCTGCTTAGTAGAGCATGCAGAAATTGATTCTACAACAAAACGTATGGAAAAGATTGCAATTCCTTT
 CCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATGCAAAGTATTGATTCTGTTGTTGAATTTTGAACGAA
 AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAAT
 TCTGAGGATTTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATCAGTCAA
 AGAATTGTGTGCTTAGTTCCTGATAAGTCTGGAAGCATGGGGGTAAAGACCGCTTAAATCGAATGAATCAAGCAGCA
 AAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGGTGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTA
 AATAAGCTAATCCAAATAAAAAGCAGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCTCTGGGAGGA
 ACTTCCATCTGCTCTGGAATTAATATGCATTTCAGGTGATTGGAGAGCTACATCCCAACTCGATGGATCCGAAGTA
 CTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGATTTGATGAAGTGAACAAAGTGGGGCCATTGTTTCA
 TTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTATGTT
 TCAGATGAAGCTCAGAACAAATGGCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAA
 TCCCTTCAGCTCGAAAGTAAGGGATTAACTGAATAGTAATGCCTGGATGAACGACACTGTCAATTTGATAGTACA
 GTGGGAAGGACACGTTCTTTCTCATCAGATGGAACAGTCTGCCTCCCAAGTATTCTCTCTGGGATCCCAAGTGAACA
 ATAATGGAAAATTTACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCCAGGAAGTCAAAAGGTGGGCACT
 TGGGCATACAATCTTCAAGCCAAAGCGAACCAGAAAACATTAACTATTACAGTAACTTCTCGAGCAGCAAAATCTTCT
 GTGCCTCCAATCAGAGTGAATGCTAAAATGAATAAGGACGTAACAGTTTCCCAGCCCAATGATTGTTTACGCAGAA
 ATTCTACAAGGATATGTACCTGTTCTTGGAGCCAAATGTGACTGCTTTCATTGAATCAGAGAATGGACATACAGAAGTT
 TTGGAACTTTGGATAATGGTGACGGCGCTGATTCTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT
 ACAGAAAATGGCAGATATAGCTTAAAAGTTGGGGCTCATGGAGGAGCAAACTGCCAGGCTAAAATACGGGCTCCA
 CTGAATAGAGCCGCGTACATACCAAGGCTGGGTAGTGAACGGGGAAATTGAAGCAAAACCCGCAAGACCTGAAATTTGAT
 GAGGATACTCAGACCACCTTGGAGGATTTAGCCGAACAGCATCCGAGGTGCATTGTTGGTATCACAGTCCCAAGC
 CTTCCCTTGCCTGACCAATACCCACCAAGTCAATCAGACCTTGATGCCACAGTTTATGAGGATAAGATTATTCTT
 ACATGGACAGCACCAGGAGATAATTTTATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT
 GATCTAAGAGACAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCAACAAAGGAGGCCAATCCCAAGGAA
 AGCTTTGCATTTAAACAGAAAAATATCTCAGAAGAAAATGCAACCCACATATTATTGGCCATTAAAAGTATAGATAAA
 AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACTTTGTTTATCCCTCAAGCAAACTCTGATGACATTGAT
 CCTACACCTACTCCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTTCTACGCTGGTATTG
 TCTGTGATTGGGTCTGTTGTAATTTGTTAACTTTATTTAAGTACCACCATTGGAACCTTAACGAAGAAAAAATCTTC
 AAGTAGACCTAGAAGAGAGTTTTTAAAAAACAAAACAATGTAAGTAAAGGATATTCTGAATCTTAAAATTCATCCCAT
 GTGTGATCATAAACTCATAAAAAATAATTTAAGATGTCGGAAGGATACTTTGATTAAATAAAAAACATCATGGATA
 TGTA AAAACTGTCAAGATTAAAAATTAATAGTTTCAATTTATTGTTATTTTATTGTAAGAAATAGTGATGAACAAAG
 ATCCTTTTTCATACTGATACCTGGTTGATATATTATTGATGCAACAGTTTTCTGAAATGATATTCAAATTCATCAA
 GAAATTAATCATCTATCTGAGTAGTCAAAATCAAGTAAAGGAGAGCAAAATAACACATTTGGAAAAAATAAAAA
 AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTASTYLE
 ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
 IHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
 RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFMQSIDSVVEFCNEKTHNQEAP
 SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLVLDKSGSMGGKDRNLNR
 MNQAAKHFLQLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
 YAFQVIGELHSQLDGSEVLLLTGDNNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
 SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTNSNAWMNDTVIIDSTVGKDTFFL
 ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
 AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
 GADSFKNQGVYSRYFTAYTENGRYSLKVRAGGANTARLKLRPPLNRAAYIPGWVNGEIEANPP
 RPEIDEDTQTTELEDFSRASGGAFFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDN
 FVDGKVQRYIIIRISASILDLRDSFDDALQVNTDLSPEANSKESFAFKPENISEENATHIFIAI
 KSIDKSNLTSKVSANIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI
 VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGAAAGACCATACGTCCCCGGGCAGGGGTGA
CAACAGGTGTCATCTTTTGTATCTCGTGTGGTGCCTTCTTATTTCAAGGAAAGACGCCAAGGTAATTTTGACCCA
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCAAGTTATGCCAGGATTTACTAGAGAGTGTCA
ACTCAACAGCAAGCGGCTCCTTCGGCTTAACCTGTGGTTGGAGGAGAGAACCCTTGTGGGGCTGCGTCTCTTAGCA
GTGCTCAGAAGTGACTTGCCTGAGGGTGGACCAGAAGAAAGGAAAGTCCCTCTTGCTGTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAAACCTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCTTTAAA
AGTAGAGAAGCTGCTCTGTGTGGTGGTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCCAAACGCATGCTTCCTGTGGTCTAGCCAGGGAAGCCCTTCCGTGGGGCCCCGGCTTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCCCTGAATGATGATGGTTCCGCGGGGGCTGCTTGCCTGGATTCCCGGGTGGTG
GTTTTGCTGGTGCCTCTGCTGTGCTATCTCTGTCTGTACATGTTGGCCTGCACCCCAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGGCCAACAGCCCCACGGGGAAGGAGGGGTACCAGGCCGTCTTCAGGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGG
AATGGGCAGTACCAGCCAGCGATGCTGCTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTC
CTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCCACAGATGTCAGCAGTG
CCTTTCCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGGCCTTACCCGCCACCCCGAGGAGAAGCCTGTG
AGGAAGGACAAGCGGGATGAGTTGGTGGAGCCATTGAATCAGCCTTGGAGACCTTGAACAATCCTGCAGAGAACAGC
CCCAATCACCGTCTTACACGGCCTCTGATTTATAGAAGGGATCTACCGAACAGAAAGGACAAAGGGACATTGTAT
GAGCTCACCTTCAAAGGGGACCACAAACAGGAATCAAACGGCTCATCTTATTCGACCATTACGCCCCATCATGAAA
GTGAAAAATGAAAAGCTCAACATGGCCAACACGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTC
CGGCAGTTCATGCAGAAATTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG
AAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAAACACTTCCAAAGCTGCCAACTTCAGGAACCTTACCTTCATC
CAGCTGAATGGAGAATTTCTCGGGGAAAGGACTTGATGTTGGAGCCGCTTCTGGAAGGGAAGCAACGCTCTCTC
TTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG
GTATTTTATCCAGTCTTTTCACTCAGTACAATCCTGGCATAATATACGCCACCATGATGCAGTCCCTCCCTTGGAA
CAGCAGCTGGTCATAAAGAAGGAACTGGATTTTGGAGAGACTTTGGATTGGGATGACGTGTCAATCGGTTCAGAC
TTCATCAATATAGGTGGGTTTGTCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC
CACAGCAACCTCATAGTGGTACGGACGCCTGTGCGAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG
CTGACCCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCCACGGCCAGCTGGGCATGTG
GTGTTCAAGGCAGCAGATAGAGGCTCACCTTCGCAACAGAAACAGAAAGACAAGTAGCAAAAAACATGAACCTCCAG
GAAGGATTGTGGGAGACACTTTTCTTCTCTTGTCAATTAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA
GGACGCAAAAAGAAATTGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTGTGGGGCTTTTACAACAGA
AATCAAAATCTCCGCTTTGCCCTGCAAAAGTAACCCAGTGTGACCCCTGTGAAGTGTCTGACAAAGGCAGAAATGCTGTG
AGATTATAAGCCTAATGGTGTGGAGGTTTGTGTTGTTTACAATACACTGAGACCTGTTGTTTGTGTGCTCATTGA
AATATTATGATTTTAAAGCAGTTTTGTAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCTCATATGAATGA
GCCATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAATATCAGAAGGCAGGAGAGGATAGGCTTATATGATACT
AGTGAGTACATTAAAGTAAAAATAAATGGACCAGAAAAAGAAACCATAAATATCGTGTATATTTTCCCAAGAT
TAACCAAAAATAATCTGCTTATCTTTTGGTTGTCCTTTTAACTGTCTCCGTTTTTTCTTTTATTTAAAAATGCAT
TTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTAC
ATTTTATATTTTAAAGAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCATTGAGACATAGGGAAGGAATGCTTTGTACT
CCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCATTGAGACATAGGGAAGGAATGCTTTGTACT
AATACAGACGTACAGATACTTTCTCTGAAGAGTATTTTGAAGAGGAGCAACTGAACACTGGAGGAAAAAGAAATGAC
ACTTTCTGCTTTACAGAAAAGGAAACTCATTGAGACTGGTGATATCGTGATGTACCTAAAAGTCAGAAACCATTTT
CTCCTCAGAAGTAGGGACCGCTTTCTTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACAATCTCTTTTC
AAAACAGGGTCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATATATATATATATTGT
GAAAGATCAATCCATCTGCCAGAACTAGTGGGATGGAAGTTTGTCTACATGTTATCCACCCAGGCCAGGTGGAAG
TAACTGAATTATTTTAAATTAAGCAGTTCTACTCAATCACCAAGATGCTTCTGAAAATTGCATTTTATACCATTT
CAAATATTTTAAAAATAAATACAGTTAATACATAGAGTGGTTTCTTCAATCATGTGAAAAATTATAGCCAGCACCAG
ATGCATGAGCTAATTATCTCTTTGAGTCCTTGTCTTCTGTTGCTCAGAGTAAACTCATTTTAAAAAGCTTCAAGAAC
ATTCAAGCTGTGGTGTGTTAAAAATGCATTGTATTGATTTGTACTGGTAGTTTATGAAATTTAATTAACACAGG
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAATATGATTGTGGATATGAA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDDEQLALPRANSPTGKEGYQAVLQEW
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVNEKLNMAN
TLINIVVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNLCRLNTQPGKKVFYPVLFSQY
NPGIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLDIKGWGGEDVHLYR
KYLHSNLIVVRTPVRLFHLWHEKRCMDLTPEQYKCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:**Signal peptide:**

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKSGPGGSEHQINSEAT
FAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLNRKS
VVFTSAQATTEA

Important features of the protein:**Signal peptide:**

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTC
 TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA
 ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA
 ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGCTCTGACAGCTCC
 AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCCTGTTTCCATGCAACAAATATACTCCAATCTGA
 AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
 ACGCTGGTGTCTACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCCTTCGTCCC
 AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
 CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTTT
 TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
 GATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTA
 ACTTTATCACCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA
 AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
 GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAG
 AAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAATACCCCCGGATAAA
 ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCTGAAGAGCAGGA
 GCTCAGTTTGAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
 TCTTGGGCCCCGAAACGTTACAGTACTCATACCCCCCTCAGCTCCAAGACTTAGACCCCCCTGGCG
 CAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA
 TCCCCAAACTGGCAGGCTGTGTATTCTTTCGTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG
 AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
 CCAGACAGGCCACCAGGAGAAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGTTATA
 TGTGCAGATGGAAAACTGATGCCAACACTTCCTTTTGCCTTTTGTTCCTGTGCAAACAAGTGAG
 TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTCTAGTGT
 CTGTGAGAATTACTTATTTCTTTTCTTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA
 GGTCTCTTAACAATGATGGTGGGCCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
 AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHANLILYGNFDRFFVPAEK
IVINFITLNISSDKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
DSEENTECTSLTQQESLSRTIPDPKTVIEYEYDVRTTDICAGPEEQELSLQEEVSTQGTLLSQ
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
 CTGGGAAGATGGCCGGCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
 CAAGCCACCCTCAGTCCCAGTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
 ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
 GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
 ATCTGGCTGAAGGTCAACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
 GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
 TCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
 CCCACCCGCTGGTCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTGTGTA
 TAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTCAAGACCTCCTAGTGCCATCCCTGC
 CCAATCTAGTGAAAACAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC
 CTCCTGCAGCTGGTGAAGGTGCCCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA
 TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA
 AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCG
 TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
 ATTCATGGTCCTGTTGGACTCTGTGCTTCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
 TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
 ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT
 TCCCTCCAGTGAAGCCCTCCGCCCTTGTTCACCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
 TTTACACCAAAGGTGACCAACTTATACTCAACTGAATAACATCAGCTCTGATCGGATCCAGCTG
 ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
 CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG
 CCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
 TTGTGGAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG
 GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTG
 CCTGTGAAAAA

FIGURE 78

MAGPWTFLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTGDALVLTASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACCTTGTTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCTTCTGGGCCTGCCCCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGAGGCCTCCTGGGATTC
ATTCCTGTTGCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCCTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTTCCTGCTCATCCCAGAGAAATCGCTCCAACCTACTACGATGCCTAC
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCCTACAGCCTGACAGGTATGTGTGAAAGAACCAGGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
TCACCTTGCTGCTCCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCCTCTTGTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYYLGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAVNNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNYSYSLTGYV

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGCTCTCCATCTCTCCCTCCTTTC
 CCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCCT
 GCCCACGCTGCTTCCTGGCCCTTCTCCGACCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
 GTTGATCTGTGGCCCTGTGCCTCCGTGTCCTTTTCGCTCTCCCTTCCTCCCGACTCCGCTCCCGG
 ACCAGCGGCCTGACCCTGGGGAAAGGATGGTTCCCGAGGTGAGGGTCTCTCCTCCTTGCTGGGA
 CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT
 CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGAGCCACAAGGCCTGATGT
 ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCTT
 GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
 TCCCTCTGGACTCCGGGCCCCACCAAAGTCCTGCCAGCACAAACGGGACCATGTACCAACACGGAG
 AGATCTTCAGTGCCCATGAGCTGTTCCCCTCCCGCCTGCCCAACCAAGTGTGTCTCTGCAGCTGC
 ACAGAGGGCCAGATCTACTGCGGCCTCACAACCTGCCCCGAACCAGGCTGCCAGCACCCCTCCC
 ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA
 GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCAGTGATGCTGGGAGAAAG
 AGAGGCCCCGGGCACCCAGCCCCCACTGGCCTCAGCGCCCCCTCTGAGCTTCATCCCTCGCCACTT
 CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCCCTGAAGGAGAAACATAAGAAAGCCT
 GTGTGCATGGCGGGAAGACGTACTCCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCCTTCGGC
 CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
 CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGCTGGGAAGTGTGCAAGATTTGCCAGAGG
 ACAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTC
 CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC
 CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG
 GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG
 GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCACGAAGTCACT
 GGAACGTCTTCTAGCCCAGACCTGGAGCTGAAGGTACGGCCAGTCCAGACAAAGTGACCAAG
 ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA
 TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTYS
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTIYPCRHPKAVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGGCCCTCCGCTCAGCGAGAGCCTCTCC
 GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCATCCGTCACCTCTCCTGTCA
 TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGGTCTGAGTC
 TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAG
 GACGCAGCATTTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCCTCAGGGG
 CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
 AAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAACATTACT
 GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
 ACAGGTGTGAGCACTGGGCTCAGTTCCTCTCATTTCATCACGGGATATGTTGATAGAGACATCCAGCTAC
 TCTGTCACTCCTCGGGCTGGTTCCTCCCGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTTGTCC
 ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA
 CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
 GAGATACCTTTTTCAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA
 TTTTTTGGCATTGTTGGACTGAAGATTTTCTCTCCAAATTCAGTGGAATAATCCAGGCGGAATCTGGACTG
 GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCCGGAACACGCAGTGGAGGTGACTCTGGATCCAG
 AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTAAACCATAGAAAAGCTCCCCAGGAGGTG
 CCTCACTCTGAGAAGAGATTACAAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAAACATTA
 CTGGGAGGTGGACCGAGGACACAATAAAAGGTGGCGCTGGGAGTGTCCGGGATGATGTGGACAGGAGGA
 AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTC
 ACATTAAATCCCCGTTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCTGGACTA
 TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGTTTG
 AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAATCCCATAGTCATCTGC
 CCAGTCACCCAGGAATCAGAGAAAGAGGCCCTTGGCAAAGGGCCTCTGCAATCCAGAGACAAGCAACAG
 TGAGTCCTCCTCACAGGCAACCACGCCCTTCTCCCGAGGGTGAAATGTAGGATGAATCACATCCCACAT
 TCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGCCAAGGTGGCTTCCA
 GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGGAGGGAAGAAGG
 CTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAG
 AACCCTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
 ATCTTATTGATGACAGAGTGATCCTAATGGTTGTTTCATTATATTACACTTTCAGTAAAAAA

FIGURE 84

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH
 LYRDGKDQPFMQMPQYQGR TKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ
 VSALGSVPLISITGYVDRDIQLLCQSSGWFPPTAKWKGPQGQDLSTDSRTNRDMHGLFDVEISL
 TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
 FQWKIQAE LDWRRKHGQAE LRDARKHAVEVT LDPETAHPKLCVSDLKTVTHRKA PQEVPHSEKRF
 TRKSVVASQS FQAGKHYWEVDGGHNKRWRVGVC RDDVD RRRKEYVTLS PDHGYWVLRLNGEHL YFT
 LNPRFISVFPRT PPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
 PIVICPVTQESEKEASWQRASAI PETSNSSESSQATT PFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT
TACCCTGGCCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
TACTTCTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAGGCACCCTGGAGTCCGGCTGCCCCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCCCTCATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTGAGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCCCTCAGAAGTTGACCATGACTGTCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCTGTGCCCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAG
CCACAGCCCTGGTCTTCCCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTTGAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTGAGGGGTTCT
AGCCTCTCAGGGGCCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCCAG
CTTCTGCCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC
TCCCTTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPLDPTTRS
SVLTLIPQPDHGTSLTCQVTFPGASVTTNKTVHLNVSYPQNLMTVFQGDGTSTVLGNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLSKATSGVTQGVVGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDGTGIE
DANAVRGASQGPLTEPWAEDSPPDQPPASARSSVGELEQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
 CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
 TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
 TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCTAGTGCATTTGA
 TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTTATCTACCAGACCTTCTGTGACATGACCTCTG
 GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
 GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
 CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
 ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG
 AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
 GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
 CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTAT
 GGCCAGCGGGAATTCAGTGCGGGATTTGTTCACTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
 CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
 GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
 GGAACATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
 TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA
 CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
 AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGLIYQKYPVKYG
EGKCWTDNGPVIPVYDFGDAQKTASYISPYGQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPPQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:**Signal peptide:**

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCTCTCAGGAGTCTCCAATGTTCAAC
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCCAGCTG
TCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAGAAACCCTTCTGATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCTTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCTCGGAAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCCTTGTAAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
 TCGCTGCTGCTTCGTGTTCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
 GTCTCCCGGCTCAGAGGACCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCCT
 CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
 GGCCCCGCTTGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCCGAACCACAGCCCCC
 CACCCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
 GCCCTGAACCTGCTCGTCACAGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
 CCAACACAATGCCACAGGCCAGGGAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
 AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
 GAGTGGGAGAAGGTAGAACGGGGCCGCCGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG
 CTCCCAGACACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG
 TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACTAC
 CATAGTGATACCCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
 TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTC
 AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
 AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
 AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGTGGG
 TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
 CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTCATCAGTGTGGACA
 GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
 GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCCTTGTCGTGTGCTGAGCATGG
 CATGAGGCTGAAGTGGAACCCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
 CAGGCCACCCCTTTCCAAAATTCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
 GCACACCCATCCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC
 CGCGTGTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
 GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACCGGGGTCTCCCGGATCTGGATGGCGC
 CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGCCGAGAGCATGTGCTGGATCTGTTC
 TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT
 TGTGTGAAGAATCGTGTCTTGGAGCAGGAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
 LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
 SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP
 AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCC
 CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGAGCTTCTT
 CTGGTTGGTGTCTCTACTGATTCGTCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA
 AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTGTCTATATCCAAGAA
 ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
 AGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA
 GTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTCTAT
 GGAGATTCTCCTCAATTCTTCCTTTATTAGCTTTTCATGACGCTGGTCATTATCTTGCTGCATGT
 ATTCTGGGGCATTGTATTTTTTGTGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC
 TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
 TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
 AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAAGTTCTTCTTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT
 TTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLPGPTVGIHGDSPOFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKKGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSCLKCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAAGTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAAACA
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTACACAACCTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTTGTTCCTGGGAGGCATCCTGCCCACCAGTCAGGCAGGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCAGCAGGCCCGCCTCCCAACTCCCAGTGGCACAGATGACGACTTTCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATCTTAATT
TACCTGAAAATATTCTTGAAATTTAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAACATATTTGGAAAAGTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
 CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCTTGACCATG
 GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGCCAGCCTGC
 AGAGCTGTCTGTGAAGTTCCAGAAAATATGGTGAAATTTCCCTTTATACCTGACCAAGTTGC
 CGCTGCCCGCTGAGGGGGCTGAAGGCCAGATCGTGCTGTGACGGGGACTCAGGCAAGGCAACTGAG
 GGCCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACAGGGCCCTGGACCGAGAGGA
 GCAGGCAGAGTACCAGCTACAGGTCACCCCTGGAGATGCAGGATGGACATGTCTTGTGGGGTCCAC
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCATTTCTCTCAAGCCATCTAC
 AGAGCTCGGCTGAGCCGGGGTACCAGGCCCTGGCATCCCCCTCCCTCTTCTTGGAGCTTCAGACCG
 GGATGAGCCAGGCACAGCCAACTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC
 CTTCCCCAGACATGTTCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
 ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA
 CCAGGCCTCAGGCCACAGGCCACTGCCACCGTGGAAGTCTCCATCATAGAGAGCACCCTGGGTGT
 CCGTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCCGCACCACATGGCCCAGGTA
 CACTGGAGTGGGGGTGATGTGCACTATCACCTGGAGAGCCATCCCCCGGGACCCCTTTGAAGTGAA
 TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCAGGCTGAGTACCTGCTCC
 AGGTGCGGGCTCAGAATTCCATGGCGAGGACTATGCGGGCCCTCTGGAGCTGCACGTGCTGGTG
 ATGGATGAGAATGACAACGTGCCATCTGCCCCCTCCCGTGACCCACAGTCAGCATCCCTGAGCT
 CAGTCCACCAGGTACTGAAGTGACTAGACTGTGACAGAGGATGCAGATGCCCCCGGCTCCCCCA
 ATTCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC
 CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT
 CTTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGAGCACGTGTGAAG
 TCGAAGTCGAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCAGATTGGGCTT
 ATAAGCCTCCCTGAGGATGTGGAGCCCGGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
 CCTCGAGCCCGCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTG
 GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
 GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCCC
 AGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGATGCCACCCCCCAAGTTGG
 ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCCAGCCGGCTCTTCTCTGCTGACCATC
 CAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT
 CTGCATTGAGAAATTTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGG
 ACACCTACACGGTGTCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
 TACCTCTGCACACCCCGCCAAAGACCATGGCTTGATCGTGAGTGAGCCAGCAAGGACCCCGATCT
 GGCCAGTGGGCACGGTCCCTACAGCTTACCCTTGGTCCCAACCCACGGTGCAACGGGATTGGC
 GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCACGT
 GAACACATAATCCCCGTGGTGGTGCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTCGAGTGAT
 CGTGTGTGCTGCAACGTGGAGGGGCAGTGATGCGCAAGGTGGGCGCATGAAGGGCATGCCCA
 CGAAGCTGTGCGCAGTGGGCATCCTTGTAGGCACCCCTGGTAGCAATAGGAATCTTCTCATCCTC
 ATTTTCACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTCCCCCT
 GAAGGCGACTGTCTCAATGGCCCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
 TCCCCCTGGGAGAGAGGCCAGCACCACAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA
 TCTGCCCTGGGTGGAGGCACCATCACCATCACAGGCATGTCTGCAGAGCCTGGACACCAACTT
 TATGGACTGCCCATGGGAGTGCTCCAAATGTGAGGTGTTTGCCCAATAATAAGCCCCAGAGAA
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAQYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKMDGQASGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFVEVNAEENLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTADADLEPAFRLMDFAIERGDTEGTFGLDWEPPDSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPI SAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWWVEP
REHIIPVVVSHNAQMWQLLVRVIVCRNVEGQCMRKVGRMKGMPTKLSAVGILVGTLVAGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCCAGCCTG
 AGTCCAAGATTCTTCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
 TCTCTTACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC
 ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
 TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCATGAGACTAGCACC
 TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
 TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
 GCATAGTCACCAACTCTGAGTTCATACAACTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
 TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
 GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCACCAACT
 CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
 AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
 CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
 GGCCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
 AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA
 GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
 TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
 GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGACAGCCACCAACT
 CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCTCC
 AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC
 CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA
 GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
 CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA
 ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
 ATGCACACAACTTCCATAGTGCACTCTACTGCAGTGGAGTGGGCAAAGCCTGGTGGGTCCCTGGT
 GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGC
 TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT
 CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCACAGGCCAG
 GTGGAGTCCCTAACTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
 ACAGCGGGCCCTAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA
 CCCAAGACCTGGTTTCTTTTCAATCATCCAGGAGACCCCTCCAGCTTTGTTTGGATCCTGAA
 AATCTTGAAGAAGGTATTCCTCACCTTTCTTGCTTTACCAGACACTGGAAAGAGAATACTATAT
 TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACAGACAAAGAGAAGCTGTGCTTG
 CCCGGGGTGGGTATCTAGCTCTGAGATGAACCTCAGTTATAGGAGAAAACCTCATGCTGGACTC
 CATCTGGCATTCAAAATCTCCACAGTAAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA



FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
 SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
 GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
 TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
 STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
 GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
 TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES
 STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
 AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLPGPGGNHGAPHRPRWSPNWFWRPVSII
 AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTACAGGTG
GGAGGCTGGAACTATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGGAATTTGGTGTCTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTCTGGAAGGGATTACTTTATCTACTCTGTTTTGGGGAAGCTTTTTGGAAGCATT
TCATGCTGAGTCCCTTTTACCTTTGATGTTTGTAAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAAGT
GATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCTATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCTCAAAGCGAGTCTCAAAGGTGTTCTGTGATTGTTGGTGGCCATGCAGGCTGCTGCCTATAT
CTTCATTATAGGAAATGGAAGGATGACAAGAGCCATTTCGAAGACATGATTGATTACTTTTGTG
ATATTCACGAACCACTTCAACTCCTCATATTTCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAAACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG
GGGAGAAGAATTTTATTTTACCGGACAGAGTGTCTTCCACCTTGCAAGTCTGAACCTCAGGGTC
CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTTCCAGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAACTTGCATGTTACCGACTTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTGCATGACTATGTGCAATATTTCTTACT
GCCATCATTATTTGTTAAAGATATTTGCACTTAATTTGTGGGAAAAATATGCTACAATTTT
TTTAATCTCTGAATGTAATTTTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVLLNPWSINEAVSSYCTYFIKQDSKSF GIMVSWKGIYFILTLFWGSFFGSIFMLSP
 FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
 FLWNCLMRYSYLRLEKICLKASLKGVPFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
 LQLLIFFEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
 YPHNIPQSEKHLQGD FPREIHFHVHRYPIDTLPTS KEDLQLWCHKRWEEKEERLRSFYQGEKNF
 YFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
 LEIIE LACYRL LHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGAÇC
 TCCAAATCATCCATCCACCCCTGCTGTCTCATCTGTTTTCATAGTGTGAGATCAACCCACAGGAATA
 TCCATGGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
 CACTGGACCGGGCAAGTTTGTCCAGGCCCTGGTGGGGGAGGACGCCGTGTCTCCTGCTCCCTCT
 TTCTTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTTCAGGAATCAGTTCCATGCTGTGGTC
 CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAAGTGA
 GTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTTAAGGCTAAAAACATCACTCCCTCGG
 ACATCGGCCTGTATGGGTGCTGGTTTCACTTCCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
 CGGGTGGCAGCACTGGGCTCACTTCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA
 GTTACTCTGCCTGTCTCAGGCTGGTTCCCCCAGCCCCACGCCAAGTGGAAAGGTCCACAAGGAC
 AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
 ATTTAGTCCAGGAAAATGCTGGGAGCATATGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
 GGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTA
 TTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGTGTCATGGGGATGATAATTGTTTTCTTC
 AAATCCAAAGGGAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAAATTGAG
 AGACGCCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG
 TTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCCTCACTCTGAGAAGAGA
 TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
 CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGGAAGAACA
 ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTT
 ACATTCAATCCCCATTTTATCAGCCTCCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCT
 GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTATACCCTGC
 TGACATGTGAGTTTGAAGGCTTGTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAG
 GGGACTCCCATATTCATATGTCCAGTGTCTGGGGATTGAGACAGAGAAGACCTGCTTAAAGGGC
 CCCACACCACAGACCAGACACAGCCAAGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTTCTCT
 CCGGAGCCTGCGCACAGAGAGTCAAGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC
 TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAG
 TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAA
 CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTCACGG
 ACAGTGATTCCTGCCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
 TGAGGGCACAGTGTGCTAATGATGTGTTTTTATATATACATTTTCCACCATAAACTCTGTT
 TGCTTATTCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC
 TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTTCACTGATTCTATAAGCCCAGCAT
 TACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC
 CTCATTAACACAGACACAAAAATTTAAATAAAATTTAACAATTAAACTAAACAATATATTTA
 AAGATGATATATAACTACTCAGTGTGGTTTGTCCCAAAATGCAGAGTTGGTTAATATTTAAAT
 ATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAGAAAACCATAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDA^vFSCSLFPETSAEAMEVRFFRNQFHAVVH
 LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
 VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGFQGGDLSSDSRANADGYSLYDVEISI
 IVQENAGSILCSIHAEQSHEVESKVLIGETFFQSPWRLASILLGLLCGALCGVVMGMIIVFFK
 SKGKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
 TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
 FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
 TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTNYNSTLSFTTDKLY
 AEFGREASNNFTEMSQRLESMVKNAFYKSPLEEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
 PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
 TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
 RRIIVHEKYKHPSHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
 QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLVSSDARDIWYLAG
 IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCTG
 CCCCATGAGCCCCCGCGTCCCGACTATCCCAGGCGGGCTGGGGCACCAGGCCCCAGC
 GCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT
 ACGGGGCTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA
 TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
 ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTGACTTACTTTCT
 GTGGTTCATCCGTCATGATTGCTGTTTGTCTGTTTCTTATCATTGTGGGGATGTTAGGATATTG
 TGGAACGGTGAAAAGAAATCTGTTGCTTCTGTCATGGTACTTTGGAAGTTTGCTTGTCTTTTCT
 GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTATGGTTCCAGTACAATGGTCA
 GATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
 TGCTTGAATTTTTTTCAGAGAGAGTTTAAAGTGTGTGGAGTAGTATATTTCACTGACTGGTTGG
 AAATGACAGAGATGGACTGGCCCCAGATTCCTGCTGTGTAGAGAATCCCAGGATGTTCCAAA
 CAGGCCCACCAGGAAGATCTCAGTGACCTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCTCTT
 TTTGAGAGGAACCAACAACACTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA
 TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
 ACAGACCAAAATGATGTCCTTGAAGAATGACAACTCTCAGCACCTGTCATGTCCCTCAGTAGAACT
 GTTGAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT
 TTGAGATGGAGGAGTTATAAAAAGAAATGTCACAGAAGAAAACCAAACTTGTTTTATTGGACT
 TGTGAATTTTTGAGTACATACATATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA
 TAACACCTAAGCATATACTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCTAAGTC
 ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
 TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTCCGCA
 TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
 CTAGTATATAAGTACTAATTAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
 CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACT
 ACCTAAATGTGATTTTTGCTGGTTACTAAAATATCTTACCCTTAAAGAGCAAGCTAACACAT
 TGTCTTAAGCTGATCAGGGATTTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAGTCGAT
 TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT
 ATTTTTAGCCTTTCCGTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
 TGTTATTTAAATACTTAACCACTAATTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTC
 AGAATGTAGTCTGGTCTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTTCAGA
 AAGGACTTGTATGCTGTTTTTCTCCCAAATGAAGACTCTTTTGGACACTAAACACTTTTTAAAAA
 GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTCTACAGAAAA
 TAGTGTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAACATGTGACAATTTAGAGATT
 CTTTGTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTTTTTACAA
 GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCAATTTTACTGTATTTTGTGTATTTTGTGTTAT
 TTCTCAGAATATGGAAGAAAATTTAAATGTGTCAATAAATATTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP
 VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
 LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ
 EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPGTDQM
 MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCCTCATCCTCTTCTCTGATAAAGCCCCCTACCAGTGCT
 GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCCTAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA
 CCAGTATTAAGAGGATTTTCCAGTGTTTCTGGCAGTTGGTCCAGAAGGATGCCTCCATTCTGCTTCTCACCTG
 CCTCTTCATCACAGGCACCTCCGTGTACCCGTGGCCCTAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGC
 CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCTCTCTATGTGACAACCATGTGAATGGGGAG
 TGGTACCCTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
 CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAAGGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
 TCAATGGGAAGTGTGTCTCTGGAACACCACGGTGGAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
 CTGACCAAGCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCGACGAGGACTGCCATGG
 CAGCTGCTCAGATACCAGCGAGTGACATGCGCTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
 ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTACCCTGTG
 GAGTGTGGGGTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
 TGGTGGCTGCAGCCACTCTTGCCTTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGGCCTGGTGTGT
 CTGAGGATAACCACACTTGCCAAAGTCCCTGTGTTGTGCAAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
 CTGTTGGTGGCCTGGAGCTCTCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT
 CCTCTTCTCTCAAGACATGTGGTACAGTGGTGCATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
 CAGGTCTACCCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG
 ACCTGCGAGTTTCCACGCTGTACACCATTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAAT
 CATGAGCCGAAATCATGGGATCTTCCCATTCACCTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC
 GGGAAAGCTCTGCCACCCCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
 TTGGAAGCTTGGTGGAGAGCTGCTTGGCACCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCAT
 CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAAGCACTTCCAGG
 TCCCTGTCTTCAAGTTTGTGGGCAAGACCACAAGGAAGTGTCTGCACTGCCGGGTCTTGTCTGTGGAGTG
 TTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGCGAATGCGTCTGGGGCAGGAGGAGAGGACTCAGC
 CGGTCTACAGGGCCAGACGCTAACAGGCGGCCGATCCGCATCGACTGGGAGGACTAGTTCGTAGCCATACCTC
 GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACC GCCCTTAAGAACATCTGCCAACAGC
 TGGGTTCAGACTTCACACTGTGAGTTCAGACTCCAGCACCACTCACTCTGATTCTGGTCCATTCACTGGGCA
 CAGGTCACAGCACTGCTGAACAATGTGGCCTGGGTGGGTTTCATCTTTCTAGGTTGAAACTAACTGTCCA
 CCCAGAAAGCACTCACCCATTTCCTCATTCTTTCTTACACTTAAATACCTCGTGTATGGTGCAATCAGAC
 CAAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCTAGAAAAATTAAACAGTTACTGAAATTATGA
 CTTAAATACCCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTCAATTCAATGCAGACTAA
 TTATAGGGAATTTGGAAGTGTATCAATAAACAGTATATAATTTT

FIGURE 110

MPFFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
 TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV
 YCGHFYDIDCEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
 LRSDGKTCEDEVEGCHNNNGGCSHCLGSEKGYQCECPRGLVLSEDNHTCQVPVLCKSNAIEVNIIPRELVGG
 LELFLTNTSCRGVSNHVNILFSLKTCGTVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
 CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEOPYREALPTLKLKRLDSLYFGIEPVVHV
 SGLESLVESCFAPTPTSKIDEVLKYLLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDKHEVFLHCRV
 LVCGVLDERSRCAQGGCHRRMRRGAGGEDSAGLQGQTLTGGPPIRIDWED

Important features of the protein:**Signal peptide:**

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
 522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
 GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
 CCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
 TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCCGAAACCCCGTATCCCC
 ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
 GGTGTCTCATCAAGGTGATTCTGGATAAACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
 GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
 CCCGAAGGGCCTGCAGTGGCAGTCCGCCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
 AGGGAACCTGGTTCTCTGCCTGTTTCGACAACCTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
 GCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
 GAGCTTCGCATGCGGAACTCAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCAGTGTCTTGCCTG
 TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGGAGGAGGCCCTGTGTGGATTCTTGGCCTTGGCAGG
 TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA
 GCCCCTGCTTCAGGAAACATAACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAG
 CTCCCCATCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCACCCCATGTACCCCAAAGACAATGACA
 TCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT
 GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAA
 GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCGT
 ACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGT
 GACAGTGGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
 CTGCGGGGGCCCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
 GGAAGGCTGAGCTGTAATGCTGCTGCCCCCTTGCAGTGCTGGGAGCCGCTTCCTTCTGCCCCGCCCCACCT
 GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCCTGGGTACACCCCTCTGCCCACAGCCTCAGCAT
 TTCTTGGAGCAGCAAAGGGCCTCAATTCCTGTAAGAGACCCTCGCAGCCCAGAGCGCCCAGAGGAAGTCA
 GCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCT
 CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAAAGCC
 CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA
 GCCTACTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
 ACTGTTGTCATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA
 AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIALLSLASIIIVVVLIKVILDKYYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGEEASVDSWPWQVSIQYDKQHVCSSILDPHWLTAHCFRKHTDVFNWKVRAGSDKL
GSFPSLAVAKIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCQGDSGGPLMYQS
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCAAGGTGCTGTGATTA
 TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTTTCAGCAACTAAAAAGCCACAGGAGT
 TGAAGTGTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAAAATC
 TGTTTTTTGTTCTCTTGTAAGTAGCCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT
 GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCTCGAAG
 CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT
 GGAGGCACTGACTCGGGCAGTGACAGGTAGCTGAGCCTCTTGCTAGCTGCGGCTTTCAAGGTGGGC
 CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTTCATAGGCGATGGCTCCCACTGCCC
 AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCGGTGGACACCTGTCTCA
 GAAGCAGTGGGTGAGACATCACGCTGCCCCCATCTAACCTTTTCATGTCCTGCACATCACCTG
 ATCCATGGGCTAATCTGAAGTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
 CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCATTCTCCTGGAACATGAGG
 GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACCTTGTCCAAATTATGGGTGAGAAAGATG
 GAGGTGTTGGGTTATCACAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGGAAGGGCTG
 CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT
 CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
 TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCC
 CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
 GGCTAATGGCTCAGTGTGGCCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT
 GTGCGAACACGGAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCAGGCAAAGGACTGTGTGG
 CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTT
 CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTGAATTTGAAACCCCAAATCCA
 AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
 TTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGG
 TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
 CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
 GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTA
 TGGTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:
amino acids 1-15

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWPAPKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA
 GGGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCTCCTGTCCATCCTGGGGCT
 GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAACCCCGTCACCT
 CCGTGTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCC
 TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCCT
 GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
 CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGCTTTTGTGAATTGCTGGAGTG
 TCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
 GATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGCGGCTCTGTTCTGTTGGGCTGGGTCTGGAGGCC
 TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAGAAGAAACCAACTACAAA
 GCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTACAAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT
 TGGGTCCAACACCAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC
 CTTCCAAGCAGACTATGTGTAATGTCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
 CCCAAAAACAAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
 CGATTTTCATCTTTGGAGAGGCCAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCACCATAAAACA
 GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACT
 TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC
 TTCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACCTTTGAAAGGAAA
 GAGTAGACCCAAAGATGTTATTTTCTGCTGTTGAATTTTGTCTCCCCACCCCAACTTGGCTAGTAATAA
 ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT
 ACACTGTGATCTTAAAGTTACCAAACCAAGTCATTTTCAGTTTGAGGCAACCAACCTTCTACTGCTG
 TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTCTGT
 CGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTAATTTAAGTCCTAAATATAGTTAA
 AATAAATAATGTTTAGTAAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
 GAAATGAAAAAATAATTGCTTTGACATTGCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTC
 ATGAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT
 TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCA
 TGGTGGCATAACCTGTAGTCCCAGCATTCCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGAGGT
 TGGGGTGCAGTGAGCCATGATCACACCACTGCCTCCAGCCAGGTGACATAGCGAGATCTGTCTAAAAA
 AATAAAAAATAAATAATGGAACACAGCAAGTCTAGGAAGTAGGTTAAACTAATTCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCGGAGTCC
AGCTGGCTAAAACTCATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTCTTG
GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACGTGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTTTGAAAACTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTTCCGTGATGTCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGCTCATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA
CTGGTGCTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTGTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTTCCCATCGCACAACCCAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTTTAACTTTACTATAAAGCCATGCAAAATG
ACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAAGAACTTTGATTTACTGTTCTTAAGTGCCT
AATCTTAATTACAGGAACGTGCATCAGCTATTTATGATTCTATAAGCTATTTACAGCAGAATGAGATA
TTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
CTCTTTTTATCATTACTTCAAAATGACATTGCTAAAGACTGCATTATTTACTACTGTAATTTCTCC
ACGACATAGCATTATGTACATAGATGAGTGAACATTTATATCTCACATAGAGACATGCTTATATGGT
TTTATTTAAAATGAAATGCCAGTCCATTACACTGAATAAATAGAAGTCAACTATTGCTTTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCTCCA
TTTATAATGAAGATTAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT
CTTGTGTATTAATTAACATTTTTAAAACGCAGATATTTGTCAAGGGGCTTGCATTCAAAGTGCCT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG
AAAATATTTTTGTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGAT
ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAGCAGAAAAATA
TGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAGTTGTCTTTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCAAAATTGTCATTTTGTCTGTGAAAAATAAATTTCTTCTGTGA
CCATTTCTGTTTAGTTTTACTAAAATCTGTAAATACTGTATTTTCTGTTTATTCCAAATTTGATGAA
ACTGACAATCCAATTTGAAAGTTTGTGTGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTTGCTT
TATACATTTATATTAATAAATGTACATTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFW EGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG
CCAGACGCTGACCACGTCCTCTCCTCGGTCTCCTCCGCCTCCAGTCCGCGCTGCCCCGGCAGCC
GGGAGCCATGCAGCCCCAGGGCCCCGCCCTCCCCGAGCGGCTCCGCGGCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCCGCGCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCGGGATG
GATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCTGGACACCCAACACTAC
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGATACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTCAAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTGAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTGAGGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGTGGATTAGTGATGTGCTATCTGGG
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATAATGCTTTAATTTTCATTTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG
CTTCAATCAAAGTGTTTCAATATTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTCTTTCTCTTAGTATAGCATTTTTA
AAAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAQVPGR
DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

C G T G A G C G T G T G C G C G T A C G G G G C T C T C T G C C T T C T G G G C T C C A A C G C A G C T C T G T G G C T G A A
 C T G G G T G C T C A T C A C G G G A A C T G C T G G G C T A T G G A A T A C A G A T G T G G C A G C T C A G G T A G C C C C A A
 A T T G C C T G G A A G A A T A C A T C A T G T T T T C G A A A G A A A A T T G T A G G A T C C A G T T T T T T T T A
 A C C G C C C C T C C C C A C C C C C A A A A A A A C T G T A A G A T G C A A A A C G T A A T A T C C A T G A A G A T C C
 T A T T A C C T A G G A A G A T T T T G A T G T T T T G C T G C G A A T G C G G T G T T G G G A T T A T T T G T T C T T G G A G
 T G T T C T C G G T G G C T G G C A A G A A A A T A T G T T C C A A A A T C G G T C C A T T C C C A A G G G T C C A A T T T T
 T C T T C C T G G G T G T C A G C G A G C C T G A C T C A C T A C A G T G C A G T G A C A G G G G C T G T C A T G C A A C T
 G C C C T A A G C C A A G C A A A A G A C C T A A G G A C G A C C T T T G A A C A A T A C A A G G A T G G G T T T C A A T G
 T A A T T A G G C T A C T G A G C G G A T C A G C T G T A G C A C T G G T T A T A G C C C C A C T G T C T T A C T G A C A A T G
 C T T T C T T C T G C C A A C G A G G A T G C C C T A A G G G C T G T A G G T G T G A A G G C A A A A T G G T A T A T T G T G A
 A T C T C A G A A A T T A C A G G A T A C C C T C A A G T A T A T C T G C T G G T T G C T T A G G T T T G T C C C T T C G C T
 A T A A C A G C C T T C A A A A C T T A A G T A T A C A A T C A A T T A A A G G C T C A A C C A G C T C A C C T G C C T T A T C
 C T T G A C C A T A A C C A T A T C A G C A A T A T T G A C G A A A A T G C T T T T A A T G G A A T A C G C A G A C T C A A A G A
 G C T G A T T C T T A G T T C C A A T A G A A T C T C C T A T T T T C T T A A C A A T A C C T T C A G A C C T G T G A C A A A T
 T A C G G A A C T T G G A T C T G T C C T A T A A T C A G C T G C A T T C T C T G G G A T C T G A A C A G T T T C G G G C T T G
 C G G A A G C T G C A G T T T A C A T T A C G G T T A A C C C T G A G A A C C A T C C C T G T G C G A A T A T T C C A
 A G A C T G C C G C A A C C T G G A A C T T T T G G A C C T G G G A T A T A A C C G G A T C C G A A G T T T A G C C A G G A A T G
 T C T T T G C T G G C A T C A G A C T C A A A G A A C T T C A C C T G G A G C A C A A T C A A T T T T C A A G C T C A A C
 C T G G C C C T T T T C C A A G G T T G G T C A G C C T T C A G A A C C T T T A C T G C A T G G A A T A A A A T A A A T C A G T G
 C A T A G G A C A G C A A T G T C C T G G A C C T G A G C T C C T T A C A A A G G C T G A T T T A T C A G C C A A T G A G A
 T C G A A G C T T T C A G T G G A C C C A G T G T T T T C C A G T G T G T C C C G A A T C T G A C A G C C C T C A A C C T G G A T
 T C C A A C A A G C T C A C A T T A T T A T T G G T C A A G A G A T T T T G G A T T C T T G G A T A T C C C T C A A T G A C A T C A G
 T C T T G C T G G G A A T A T A T T G G A A T C A G C A G A A A T A T T G C T C C C T T G A A A C T G G C T G A A A A G T T
 T T A A A G G T T A A G G G A G A A T A C A A T C A T T G T G C A G T C C C A A A G A G C T G C A A G G A G T A A T G T G
 A T C G A T G C A G T G A G A A C T A C A G C A T C T G T G G C A A A A G T A C T A C A G A G A G T T T G A T C T G C C C A G
 G C C T C T C C C A A A G C C G A C G T T T A A G C C A A G C T C C C A G G C C G A A G C A T G A G A G C A A A C C C C C T T
 T G C C C C G A C G G T G G G A G C C A C A G A G C C G G C C A A G A C C G A T G C T G A C G C C G A G C A C A T C T C T
 T T C C A T A A A A C A T C A G C G G G C A G C G T G C G C T T T C C T G T C C G T C C T C G T C A T C C T G C T G G T T A T
 C T A C G T G T C A T G G A A C G G T A C C C T G C G A G C A T G A A G C A G C T G C A G A C G C T C C C T C A T G C G A A
 G G C A C A G G A A A A A G A A A G A C A G T C C C T A A A G C A A A T G A C T C C C A G C A C C C A G G A A T T T T A T G T A
 G A T T A T A A A C C C A C A C A C G G A C C A G C G A G A T G C T G C T G A A T G G G A C G G G A C C C T G C A C C T A
 T A C A A A T C G G G C T C C A G G G A G T G T G A G G T A T G A A C C A T T G T A A A A A G A G C T T T A A A A G C T
 G G G A A T A A G T G G T G C T T T A T T G A A C T C T G G T G A C T A T A G G G A A C C G A T G C C C C C C T C C C C
 T T C C C T C T C C C T C T C A C T T T G G T G G C A A G A T C C T T C C T T G T C C G T T T T A G T G C A T T C A T A A T A C T
 G G T C A T T T T C C T C T A C A C A T A A T C A A C C A T T G A A A T T T A A T A C C A C A A T C A A T G T G A A G C T T
 G A A C T C C G G T T T A A T A T A A T A C C T A T T G T A T A A G A C C C T T T A C T G A T T C C A T T A A T G T C G C A T T
 G T T T A A G A T A A A A C T T C T T T C A T A G G T A A A A A A A A A A

FIGURE 124

MGFNVIRLLSGSAVALVIAPTPLLTMSSAERGCPKGCRCGKMVYCESQKLQEIPSSISAGCLG
LSLRNSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLSNSLRITIPVRIFQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQNLQWNKISVIGQTMSWTWSSLQRLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIQEILDSWISLNDISLAGNIWECNRNCSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILVIYVSWKRYPASKMLQQR
SLMRRHRKKKQSLKQMPSTQEFYVDYKPTNTETSEMLLNGTGPCCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMOMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

CGCGCGCGCGCGCGCAGGAGG

FIGURE 128

MARPGMERWRDRLALVTGASGGIGA A VARALVQQGLKVVG CARTVGNIEELAAECKSAGYPGTLI
 PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFN VNV LALSICTR
 EAYQSMKERNVDDGHIININMSGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQT HIRATC
 ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
 199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AAC TTCTAC ATGGGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTTCAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCTCCCAGGCAGTCTGCTCAGAGGGCCTCGGCCAGAAATCCAGTTCTGGTTTCATGC
CAGCCTGTAAAAGGCCATGGAAC TTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCA TTTCAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAA TGAGGAACTGAGAAAATTAATTTCTCATGTATTTTTCTCATTATTTA
TTAATTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTTATTCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCAACTAATTTTTGTATTTT TAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTATGATTACA
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTTCTTTGTGTTGGGAAC TTTGAAATTAT
ACAATGAATTATTGTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCATGTG
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTTCTGTTCTTAATTTCAATTAATAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGHGTLGESPMFPKRVCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
 ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
 TCCATCTGTGTGCGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
 GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCTTCAGAT
 TTGAAAACTTGCTGAAAGTAGAAGAATATACCTATACCACAACAGTTTAGATGAATTTCTTACCAACCT
 CCCAAAGTATGTAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
 AAATTCCTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTGTCAGTTAGCATAGAAGAGGGAGCA
 TTCCGAGACAGCAACTATCTCCGACTGCTTTTCTGTCCCGTAATCACCTTAGCACAAATCCCTGGGGTTT
 GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTCATCACCATCTCTTCAAG
 GTCTCACTAGTCTAAAACGCTGGTTCTAGATGGAAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
 TTCTTCAACCTAGTTAATTGACAGAGCTGTCCCTGGTGCGGAATTCCTGACTGCTGCACCAGTAACCTT
 TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCAAATGCTTTTT
 CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
 GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
 GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
 AGGTTCTGGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGC
 ACCATTGAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCAAGGACAGTGGCCAGCTCCAGTGAC
 CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA
 CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTATCTCTTGGAACCTTGCTCTACCTATG
 ACTGCTTTGAGACTCAGCTGGCTTAAACTGGGCCATAGCCCGCATTTGGATCTATAACAGAAACAATTGT
 AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTT
 CCATGGAAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAACTGCACCCCTT
 CGAATGTACAACCCTACAACCACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC
 TTTGGCTGCCATCATTTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTGGTATG
 TTCATAGGAATGGATCGCTCTTCTCAAGGAACGTGTCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
 GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAACTTCTTTTTCAGATGTTACCAAT
 AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAATACACACCATATTTTCCTCCTAATGGAATGAATCTGT
 ACAAAAACAATCACAGTGAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
 TCACACTCATTGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATG
 GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAGFIYCNDRFLTSIPTGIPEDATTLYL
 QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
 EELHLDNDSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
 QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
 RVPPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDL DNITQLILRNNPWYCGCKMKWVRDWLQSLPV
 KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
 IKNPKLTKDQQTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
 IVTGERSEYLVTALEPDSPIKVMVPMETSNLYLFDETPVCJETETAPLRMYNPTTTLNREQEKE
 PYKNPNLPLAAIIGGAVALVTIALALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
 ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLKYNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645**Amidation site.**

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCAGGCCACCCAGGCTTCTTGGCA
 GCCCTGCCGGGCCACTTGTCTTCATGCTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
 TGCAGAGGCAGTCTGGGCTTGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
 GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCCTGTGGT
 CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
 GCGCTGGAGAGAGCCACCGTCTTCCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
 GGGGGTCCGAGTGTCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGCCCAGGAGCCCCCTGC
 TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
 CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCTCCAGCCCCG
 GTTTTGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCTGGGC
 CCCAGGACTCATTTCTCAGAGGAGAGAAGTGACGTGTGCCGTGGTGCAGCTGCTGGGAACCGGGACG
 GACAGCAGCGAGCCCTGCGGCCCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
 AGGCTACTGCCTGTCCACCAACTGCTCTTCTTCTCTGCGCCAGAATGAGGGGATGCACACAGG
 GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTGAACCGC
 AGAGCTGAGGCCATCGGATACGCCTACCCTACCCGGGACATCTTCATGGAAAACATCATGTTCTG
 TGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
 AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
 TATCAGCAGCATTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT
 TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC
 AATTCTCTTGCTCATCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT
 TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
 CTCAAGAGATCCGCCCACCTCAGGCTCCCAAAGTGTGGGATTATAGGTTGTGAGCCACCGTGTCTG
 GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
 ATGGGGGCTCTCTCCCCTAGATGGCTGCTCCTCCCAACACAGCCACAGCAGTGGCAGCCCTGG
 GTGGCTTCTTATACATCCTGGCAGAATACCCCCCAGCAAACAGAGGCCACACCCATCCACACCG
 CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
 TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
 ATAAGCAAAGCCACCCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
 GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLARAQGAERVTSSSEQRPAMASLGLLLLLLLLTALPPIWSSSLPGLD
 TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
 GMLGEKLEAAIQRSIHLYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
 RSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
 YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSQKQEGCGF
 EPDAEDEELSKAIQYQOHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCILIP
 SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEAGSGSGVGIGDRFKIEGRAVVPGVKPDWISAA
RVLVDGEEHVGFLLKTDGSEFVVDIPSGSYVVEVVS PAYRFDPPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLI FVLLPKVVNTSDPDMRREME
QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGGGA
 CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAAACAA
 GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
 CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
 CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAGGCTTGGCAGT
 TTTTCTTACTCCTGTGGTCTCCAGATTTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
 CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
 AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
 TGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
 AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA
 AACTACCAGACCCCTGACCATTTACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA
 CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
 AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
 TTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAGGAGACAGAAATAGGAGGAAAGTGATGCTG
 CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
 ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
 CTTCTTGCATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
 TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTGTATTTA
 ATGTATTTATTTTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATATTTATAACCTG
 ACTAGAGCAGGTGATGTATTTTTATACAGTAAAAAAAACCTTGTAATTTCTAGAAGAGTGG
 CTAGGGGGGTTATTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
 ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTGTGTGGAATTGCAC
 ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG
 CCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTCTTTGCATACCAAAAAAAA
 AAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPG LTCWALTAEPGWGQNKGATTCATNSHSDSEL RPEIF
 SSREAWQFFLLLWSPDFRPKM KASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
 FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
 SLANSFLT IKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
 E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCGGCATCCAGCC
TAGCGTGTCCACGATCGGGCTGGGGCTCCGGGACTTTCCGTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCCTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCAGCGCCGAACCCCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC
ACCACCTCTCTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTTG
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTTACCTTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTCTGCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA
TTCCCAAAGCATTGTGTGGAATATGATGGAACAACCTCATTTTCGTGTCAGATTACACAGAGGT
GGATAATAATGTACAGAGGCATTTGGATAAAGTATTAAAAAGAGGAGATTGGGACATATTAAATCC
TCCACTACCTGGGGCTGGACCACATTGGCCACATTTAGGGCCCCAACAGCCCCCTGATTGGGCAG
AAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
GACGCCTTTACCCAATTTGCTGGTTCCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG
GGGCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTTAAATCAGTTCTGCGTTTGAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCCAAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCCTATTCCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAACTGTTGCAAGAG
AATGTGCCGTATATGAAAAAGATCCTGGGTTTGAAGCAGTTTAAATGTGAGAAAGATTGCATGG
GAACGGATCAGACTGTACTTGGAGGAAAAGCATTCAGAAGTCTTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGACAAGTGGCCAG
TTCTCACCCCTGCTCCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA
CTGTATCTCCTGGGTTTCTCTGCTCTTTTATTGGTGATCCTGGTTCTTTTCGGCCGTTTACGT
CATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCTCTCGTGGCTGGCGGCAGGCT
GCCTTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGCGAGTGGCCTGGAC
AGGGGGCTCAGGGAAGGACGTGGAGCAGCCTTATCCCAGGCCCTGCGGTGTCGCGACACAGGTG
TTCACATCTGTGCTGTGAGGTGAGTGCCTCAGTTCTTGGAAAGCTAGGTTCTGCGACTGTAC
CAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
TCGGACAGCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCCAAGATGCATCCT
GCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTCTGTTGGAATTTAGTCTTGGCC
TCGGACACCTTATTCTGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGCGGATGGTCAC
ACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCC
ACCCCAACCTGCACAGCCCTCATCCCTCTTGGCTTGAGCCGTGAGAGGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTTGTATCAGGGCACAGGCTTCTCGGAGCCAGGATGATCTGTG
CCACGCTTGCACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLEF
SKVVIVLIDALRDDVFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAACT
TCCCTATAGAAAACAACCTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCAGATAAAAACTACATACGCCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCCAAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAGTTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTGAGCGATTAGGAAACTGCCCCATTGAACGCCTTCCTCGCTA
ATTTGAACTAATTGTATAAAAACACCAAACCTGCTCACT

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTAA
TCCAGGATCCTGTCCTTCCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACCTCTGTGGG
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGTCCTGGGCGCTGCTGCTGCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAAATTTGAAGACGACATTGACAACTGCCATTTCCAAGAAAGCACAG
AGCTGAACAATACTTTCACCTGCTTCTTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTTGAGTGAAACCCACTCACAGGCTTGTCCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGGACTTGTTTGTATCCTATTTTGCATGTGTTTGAGATCTCAGATCAGTGTTTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAACATCAGCATTTTAAGAAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGAACCGGAAGGCCCTGCCCAGCACTGCCCAGATCA
CTGAGGCCCAGGTGGCTGAGAACCGCCCGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTACACGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTCACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTCTGAAGGTGCCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTCTTAGAGGTAG
GTGTATTCCCGTTTTACAGATAAGGAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCTGGCTTGCTAACCCAGGTTTTCTGCTCT
GTCCAATTCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

$$\begin{array}{ccccccc}
\left(\begin{array}{c|c} & \\ \hline & \end{array} \right) & \left(\begin{array}{c|c} & \\ \hline & \end{array} \right) & \left(\begin{array}{c|c} & \\ \hline & \end{array} \right) & \left(\begin{array}{c|c} & \\ \hline & \end{array} \right) & \left(\begin{array}{c|c} & \\ \hline & \end{array} \right) & \left(\begin{array}{c|c} & \\ \hline & \end{array} \right) & \left(\begin{array}{c|c} & \\ \hline & \end{array} \right) \\
\text{II} & \text{III} & \text{IV} & \text{V} & \text{VI} & \text{VII} & \text{VIII}
\end{array}$$

Important features of the protein:

amino acids 1-26

amino acids 157-171

amino acids 98-102, 110-114

amino acids 76-83

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATT
 CCTGATGATTTATAGACTCAAAGAAAACTATGTTTCAGAAGCTCTCTTCTTCTGGCCTCCTCT
 CTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
 AAATCTTTCATTTTGCTTGTGAGTGGGGTAGGTCACTGAGTCTTAGTTTTTATTTTTTGAAATTT
 CAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCATTGATGTGAGG
 TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAAC TTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGACTTGAC
 TCCCGCGCGCCCCAACCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCAGTCC
 CGGCCCCCTCTCCCGCCCCACACCCACCCTCCTGGCTCTTCTGTTTTTACTCCTCCTTTTCATTCATA
 ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGGAGAATGGGGTT
 CCTCGGGACCGGCACTTGGAATCTGGTGTAGTGCTCCCGATTCAAGCTTTCCTCAACCTGGAGGAA
 GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
 GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTCAGAGCAACTATCTTT
 TGTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA
 TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACTG
 ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
 TCAACTAGACGGGACTCCTTTAACCCTGTAAGACATTGTCCATAAAATCGTGCAGGATTTATGAAG
 AAAATGACAGAGCCGTGTTTGACAAGATTGTTCTAACTACTTAATCTCGGCCTTATCAGAAAAGC
 CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAAATTAATCTCAAAGGAAGCCAACA
 TTATGAGGAGGATCCCAATAAGCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
 TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
 TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAACCTCA
 ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAGAAAGCAAAAGAGAAAGAAA
 CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAATATGGAACAATATCT
 CCAGAAGAAGGTGTTTCCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAAACAAGCT
 AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
 CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATCCACAAAA
 GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC
 CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGAAATAAGAAGATTATGACCTTT
 CAAAGATGAGAGACTTCATCAATAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
 GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAATGGCAAAAGATCCAGGAGTCTTTCAA
 CTGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATCTGTGATTAAAAATTTTTGACCCAAGG
 GTTATTAGAAAGTGTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTCTTCCC
 GTAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
 QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
 FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
 VLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVT PMAAIQDGLAKGENDETVSNTLTLTNGLE
 RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
 SYLENLDEMIALQTKNLEKNATDNISKLFAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
 NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDK
 EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
 220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
 341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCTGTGGAGCTCAA
 GATGGTCCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
 ATAATAACCAAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC
 GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCCGTCATCCTGGGTGTCCAGGGTGAAG
 CCAGTGCCTGTGATGTGGGGTGGGGCAGGAGCCGACTTAACACTAGAGCCAGTGAACATCATGG
 AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTACCTTCTACCGCGGGACATGGGGCTCACC
 TCCAGCTTCGAGTCTGGCTGCCATCCCGGGCTGGTTCCTGTGCACGGTGCCTGAAGCCGATCAGCC
 TGTCAGACTCACCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCCATCACAGACTTCTACTTCC
 AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGAAGGGT
 GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTCTGCTCTCAGGACCCACGCTGACTTAG
 TGGGCACCTGACCACTTTGTCTTCTGGTTCCAGTTTGGATAAATCTGAGATTTGGAGCTCAGT
 CCACGGTCCCTCCCCACTGGATGGTGTCTACTGCTGTGGAACCTTGTA AAAACCATGTGGGGTAAA
 CTGGGAATAACATGAAAAGATTCTGTGGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCCTGCT
 TAATGGTAAGTGAACAAGTGTACCCTGAGCCCCGAGGCCAACCCATCCCAAGTTGAGCCTTATA
 GGGTCAGTAGCTCTCCACATGAAGTCTGTCACTCACCACCTGTGCAGGAGAGGGAGGTGGTCATA
 GAGTCAGGGATCTATGGCCCTTGGCCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTGCTATA
 TGCTACCTTTTCTATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGTGTCAGAA
 GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAA
 GATACAATCAAAATCCCAGATGCTGGTCTCTATTTCCATGAAAAAGTGCTCATGACATATTGAGA
 AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTTAAATTAAGATACCTATTTATATATT
 TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTGAGGGTGGTGGCAGTAT
 AGGTGATTTTTCTTTAATTTCTGTTAATTTATCTGTATTTTCTAATTTTTCTACAATGAAGATGA
 ATTCCTTGTATAAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA
 TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG
 TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
 GTGGCTGGAATCTCTGGGTAAGGAACTTAAAGAACAAAAATCATCTGGTAATTTCTTCTAGAAAG
 GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA
 ATTTGTGTCCCCCTCAAATTCACATCCTTCTTGGAACTCTCAGTCTGTGAGTTTATTTGGAGATAAG
 GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCTGCTGGATGAAGGTAGACCTAAATTCATAT
 GACTGGTTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
 AAGATGAAGGCAGAGATCGGAGTTTTGCAGCCACAAGCTAAGAAAACACCAAGGATTGTGGCAACC
 ATCAGAAGCTTGGAAGAGGCAAAGAAGATTCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG
 CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGATAAATTTCCGGCTGTTTTAA
 GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT
 CTCTCTGTGTTTACATTCTGTGTGTGTCCCCCTCCACAATGTACCAAAGTTGTCTTTGTGACCAA
 TAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTATATAAAGACACTGCAGCTTC
 TACTTGAGCCCTCTCTCTCTGCCACCCACCGCCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG
 AAGCTAGCTGCCATGCTATGAGCAGGCCATATAAGAGACTTACGTGGTAAAAAATGAAGTCTCCT
 GCCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT
 AAGTTGCTCAGTTTTGGTCTAACTTGTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESISVVPNRWLDASLSPVILGVQGG
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGAWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACTTCACCCCTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTT
CAGCCTTATATGCAGGAGGTGGTGCCCTTCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATACTAACCCCTTTCCTGCTAGAAATAA
CAATTAGATGCCCCAAAGCGATTTTTTTTAACCAAAGGAAGATGGGAAGCCAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTCATTGTAAGTGGTGTTT
TATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
TGTATTTATTATTATTATAAGACTGCATTTTATTATATCATTTTATTAATATGGATTTATTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCCGACTTGTGACTGAGTGTGCAGTGCCCAGCATGTACCAGGTGAGTGCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCCTACAGGTGGTGCAT
TCTTGGAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG
GCCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCAGGACCTGTACCACGCCCGT
TGCTGTGCCCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCGGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGTGTGCGGGCCCGT
GTGATGGGGTTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCATGAAGGGCCAGGATGCCCAGATGCTTGGCCCCGTGAAGTGCTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACTTCTGCACATTTTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTCTGTCATTTCTCTCAGGAAAGGTTTCAA
GTTCTGCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCATCCCCTGCTACCCTG
GCCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCTGGTTTTATT
TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPSKGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide: ?

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAGGCCACCAA
GATTTGTGTGACGGGCAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTTTATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTCACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTGAAGAAGATGAGGAGACA
GTAGAAGTGAACCTCACAACCACTCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTCTCAGGTGTTTGGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCACAAACAGGCGTCCCTTCCCTCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
TCACACAATTTGTTACTTCACTGAATTTCTTCAAAACCATTGCAGAAGTGAGGTATCCTTGAAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAAGAAGGCA
GCAGACAAAGTCGTCTTCCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCTTTAACCTTTTCTGCA
GTGATCTAAGAAGCCAGATTCTCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPGLDLRLRVEPVTTSVATGDYSILMNVS
 WVRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
 NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKKNEETVEVNFTTTPLGNRYMALIQH
 STIIGFSQVFEPHQKKQTRASVVIPTVDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
 SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL
 QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
 FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHVKKQVVSAGKRSQACHD
 GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
 - 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
 GTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
 CTTTTTCCAAAAGCCTGAGAGTTGCCCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
 ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAAACATCGAGAGCCGCTCCACCTCCCCCTG
 GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCAGTGTAGGA
 ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA
 GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTTCCAGTTGGAGAAGGTGCT
 GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTTAAGAGGTGCATATCC
 ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPA~~M~~VKYL~~L~~LSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
Q~~R~~VSM~~S~~RNIESRSTSPWNYT~~V~~TWDPNRY~~P~~SEVVQAQCRNLGCINAQ~~G~~KEDISMNSVPIQ~~Q~~ETLVV
RRKHQ~~G~~CSVSFQLEKVLVTVGCTCVTPV~~I~~HHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG
ACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCTCTGGAGGCTGAAGAGGGATTTC
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCTTGGGGGGGGGAGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCTGTGCCCTGGTTCTTGCTGTCTT
TGGCACTGGGGCGAAGCCCAGTGGTCTTTCTCTGGAGAGGCTTGTGGGGCCTCAGGACGCTACC
CACTGCTCTCCGGGCCTCTCTGCGCCCTCTGGGACAGTGACATACTCTGCCTGCCCTGGGGACAT
CGTGCTGCTCCGGGCCCCGTGCTGGCGCCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGCCGTGCATGGGCACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
TGCTCTCTCCAGGCCCAAGTCGTGCTCTCTTCCAGGCCCTACCCTACTGCCCGCTGCGTCTCTGC
TGGAGGTGCAAGTGCTGCTGCCCTTGTGCACTTTGGTCACTCTGTGGGCTCTGTGGTATATGAC
TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTCTTATACTCAGCCCAGGTACGAGAA
GGAACCTCAACCACACACAGCAGCTGCTGCCCTGCCCTGGCTCAACGTGTCAGCAGATGGTGACA
ACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCACTTCGGCCTCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCCAAAACCCCGTGGCACAAAAACCTGACTGGACCGCAGATCATTACCTTGAA
CCACACAGACCTGGTTCCCTGCCCTCTGTATTCAAGTGTGGCTCTGGAACCTGACTCCGTAGGA
CGAATCTGCCCCCTCAGGGAGGACCCCGCGCACACCAGAACCTCTGGCAAGCCGCGGCGACTG
CGACTGCTGACCTGCAGAGCTGGCTGCTGGACGCACCGTGTCTGCTGCCCGCAGAAGCGGCACT
GTGCTGGCGGGCTCCGGGTGGGGACCCCTGCCAGCCACTGGTCCCAACGCTTCTCTGGGAGAAGC
TCACTGTGGACAAGGTTCTCGAGTTCCCATTTGCTGAAAGGCCACCCTAACCTCTGTGTTCAAGGTG
AACAGCTCGGAGAAGCTGCAGCTGCAGGAGTGCTTGTGGGCTGACTCCCTGGGGCCTCTCAAAGA
CGATGTGCTACTGTTGGAGACACGAGGCCCCAGGACAACAGATCCCTCTGTGCCCTTGGAAACCCA
GTGGCTGTACTTCACTACCCAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCCTG
CCCCATGGACAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCTACTCTTTGCCG
CTGCGCTTTCCCTCATCTCTCTTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAA
CAGGACGTCGCTCGGGGGCGGCCGCCAGGGGGCGCGGGCTCTGCTCCTTACTCAGCCGATGA
CTCGGGTTTCGAGCGCTGGTGGGCGCCCTGGCGTCCGGCCTGTGCCAGCTGCCGCTGCGCGTGG
CCGTAGACCTGTGGAGCCGTGCTGAAGTGAAGCGCGCAGGGGGCCGTGGCTTGGTTTACGCGCAG
CGGCGCCAGACCCTGCAGGAGGGCGCGTGGTGGTCTTGTCTCTCTCTCCCGGTGCGGTGGCGCT
GTGCAGCGAGTGGCTACAGGATGGGGTGTCCGGGGCCGGGGCGCACGGCCCGCACGACGCTTCC
GCGCTCGCTCAGCTGCGTGTCTGCCGACTTCTTGCAGGGCCGGGCGCGCCGCGAGCTACGTGGG
GCCTGCTTCGACAGGCTGCTCCACCCGAGCGCCGTACCCGCCCTTTTCCGCACCGTGCCGCTCTT
CACACTGCCCTCCCAACTGCCAGACTTCTTGGGGGCCCTGCAGCAGCCTCGCGCCCCGCTTCCG
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTACGCCAGCCTGGATAGCTACTTC
CATCCCCGGGGACTCCCGCGCGGGGACGCGGGGTGGGACCAGGGGCGGGACCTGGGGCGGGGA
CGGGACTTAAATAAAGGCAGACGCTGTTTTCTAAAAAA

FIGURE 162

MPVVPWFLLSLALGRSPVVLSELERLVGFPQDATHCSPGLSCLWDS DILCLPGDIVPAPGPVLAPTHLQTELV
 LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLSFQAYPTARCVLLEV
 QVPAALVQFGQSVGSGVYDCFEAALGSEVRIWSYTPRYEKEINHTQQLPALPWLNV SADGDNVHLVNLVS
 EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQN
 LWQAARLRLTLQSWLLDAPCSLPAAEALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCVQ
 VNSSEKLQEQELWADSLGPLKDDVLLLETRGPQDNRS LCALEPSGCTSLPSKASTRAARLGEYLLQDLQS
 GQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARG
 RAALLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVWFHAQRRQTLQEGGVVLLFSP
 GAVALCSEWLQDGVSGFGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRL LHFDPAVPALFRTVPVFT
 LPSQLPDFLGALQQPRAPRSGRLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
 334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
 692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
 GCTCACGCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA
 AAACATCTTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
 AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTGACGGGATCACCCGGAAGTCTGTC
 AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTACCCTGCT
 GTCAGTGGGGAGGCGGTGAGCCACCAAGATGACTGACAGGTTGAGCTCTCTGACGACACTAC
 CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGAGATGATTGTTTCATCTA
 CCCCCACGCCAATCCGTGACGGGATGGCCACCGGCTAACCTGGAAGACATCTCCATGACCTG
 TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAATGCACCTTGGAGGGAGCAGAGAGA
 ATATGAGTTCTTCGGCTGACCCCTGACACAGAGTTCTTGGCACCATCATGATTTGCGTTCCCA
 CCTGGGCCAAGGAGAGTGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
 TACTCCTTCTCCGGAGCCTTCTGTCTCTCCATGGGCTTCTCGTGCAGTACTCTGCTACCTGAG
 CTACAGATATGTACCAAGCGCGCTGCACCTCCCACTCCCTGAACGTCAGCGAGTCTGACCTT
 TCCAGCCGCTGCGCTTCATCCAGGAGCACGTCTGATCCCTGTCTTTGACCTCAGCGGCCCCAGC
 AGTCTGGCCAGCCTGTCCAGTACTCCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
 TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
 CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTTATGCCCAACGTCGCCCCGAG
 GTCGGGCCCCCATCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCTATTCTACGCCCCACA
 GGCCATCTCTAAGGTCCAGCCTTCTCTCTATGCCCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
 CCTATGGGGTATGATGGAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCTTAAA
 CACCTTAGGCCTAAAGGTCAGCTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCTT
 TTCTCTGACGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATTGCACCAGC
 CCCTGGGGATTTGCACAGACAGAACATCTGACCCAAATGTCTACACAGTGGGGAGGAAGGGACA
 CCACAGTACCTAAAGGGCCAGCTCCCCCTCTCTCTCAGTCCAGATCGAGGGCCACCCCATGTC
 CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
 TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
 GAGCAGCCACAGAACTGGATTCTCTTTTTCAGAGGCCTGGCCCTGACTGTGCAGTGGGAGTCTTG
 AGGGGAATGGGAAAGGCTTGGTGCTTCTCCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA
 ATCCCATGCCCTGCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
 AGAGGGAGTGGCATGCAGGGCCCCCTGCCATGGGTGCGCTCTCACCGGAACAAAGCAGCATGATA
 AGGACTGCAGCGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGGTGCTCGCTGAGCCCTG
 CAAGGCAGAAATGACAGTGCAAGGAGGAAATGCAGGGAAACTCCCGAGGTCCAGAGCCCCACCTC
 CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGCCCCATTCTTGCCAGTTTC
 ACAATCTAGCTCGACAGAGCATGAGCCCCCTGCCTCTTCTGTCTATTGTTCAAAGGTGGGAAGAGA
 GCCTGGAAAAGAACCAGGCTTGGAAAAGAACCAGAAGGAGGCTGGGCAGAACCAGAACAACCTGC
 ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCTGCAGCTCA
 TTCCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTTCAGCTTCATTCCTCTGATAGAACAAAGC
 GAAATGCAGGTCCACCAGGGAGGGAGACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCCCT
 ATCCTGAGAATGGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCCCTGGACGGGTACAATAACACAC
 TGTACTGATGTCACAACCTTGCAAGCTCTGCCTTGGGTTTCAGCCCATCTGGGCTCAAATTCAGC
 CTACCACTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCCGAAGCTCGGTTTCCTC
 ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
 TCTTTAAAGTGCTTAATAGTGCTGTACATGGGCAGTGCCAATAAACGGTAGCTATTTAAAAA
 AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTPVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTelyyARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCrvKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQrVLTfQPLRFIQEHVLIpVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFFFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSLAM
EESQEAKSLHQPLGICTDRtSDPNVLHSGEEGTPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGSPWGLLESIVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTvQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAAAAGTCACCCGGGCGCGGTGGCCACAACATGG
CTGCGGGCGCGGGGCTGCTCTTCTGGCTGTTCTGTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG
TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACTTTGGGCTGGA
AGTGTGTAACACAGTTTTTGATATTTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCTGTGCCTGAGCCCGAGGCATTTCAGAGCTGATTTCAGAGGATGGAGAAGGTGCTTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGAAGCGAAAGCAGAAGTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCTGAAAACAGAAATGAGTCAGAGAGGAAGTGACAGTGCCTTATT
CATTACAGCAAAGGATTTTCGTTGGCATCAAATCTAAGTTTGTGTTTACAAAGATTGTTTTTAGTA
CTAAGCTGCCTTGGCAGTTTGCATTTTTGAGCCAAACAAAATATATTATTTCCCTTCTAAGTA
AAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFLGALWWVPGQSDLHGRRFSDLKVCGDEEC SMLMYRGKALEDFTGPD CRFVN
FKKGDDVYVYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPADETDFVCFEGGRD
DFNSYNVEELLGSLELEDSVPESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRA DSEDEGEA
FSESTEG LQGQPSAQESH PHTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNL SLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTGTCAGAGGCCGGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGGGCACAGG
CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTGGGCGAGCTGGGCTCGGGC
GGCGGGAGTAGGGCCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG
GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
TCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGAGCCTTCTGCCGCCGCGTGGTCAGC
GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAGT
GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCCTGGCTTGTGAGAGTGAGGGAGGAGTCCCTCTCA
GCCTTGAGAAATGAAGCAGAACAGAAGTAAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCTGG
TGCTGCCCAGATCTCTACCAAGTGGTCTGATGGAAGCAATTCAGTACCGAACTGGTACACAG
TGCAAGCTTCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCACCAACTGCCAATCCCTGGC
CTTGGGGGCTCCCTACCTTTACCAAGTGAATGATGACAGGTGTAACATGAAGCACAATTATATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACACCCATCAGAAATGTGGTTGTTACTGAAGCAGGTATAATTCCTCAATCTAATTTATGTTGT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGGAACTGTTGTTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAACAACTAGTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTTCCAGAATTTGTAATCT
GGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGGAAATGGCTTGAAATCACAAAGGATCTGC
AAGATGAACGTGAAGCTCCCCCTTGAGGCAAATATAAAGTAATTTTATATGTCTATTATTTCA
TTTAAAGAAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGTAAAGGATGCACCCAA
ACTTCAAACTTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTTAGAAGCAATTCCTTTTATTTCTTTTACCTTTCATAAGTTGTTATCTAGTCAATGTAA
TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGTTGATAAAA
ATGAAGTGTCTAATATTTATTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGTATTAAAG
AACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA
AATAAGAAGCTATTTTCAATTAAGTGTGATATAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
TGCTAATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTAGCTTAAATTAACAGATT
TTGTAATAATGTAACCTTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAG
TGACATACACAATATAAATCATATGTCTTACACGTTGCCTATATAATGAGAAGCAGCTCTCTGA
GGGTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAACAAAGATGGTTGTTCCGGGGTTTGGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTAGCCTCTG
ACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACATCAGTGTGGAG
ACAAGCACAGCACACAGACATTTTAGGAAGGAAGGAACACGAAATCGTGTGAAATGGGTTGG
AAGCCATCAGTGATCGCATATTCATTGATGAGGGTTGCTTGAGATAGAAAATGGTGGCTCCTTT
CTGCTTATCTCCTAGTTTCTTCAATGCTTACGCCCTTGTCTTCTCAAGAGAAAGTTGTAAGTCT
CTGGCTTTCATATGTCCCTGTGCTCCTTTTAAACAAATAAAGAGTTCTTGTTTCTGGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFCRRVVSQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSCKGRKTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:**Signal peptide:**

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-
145, 212-217